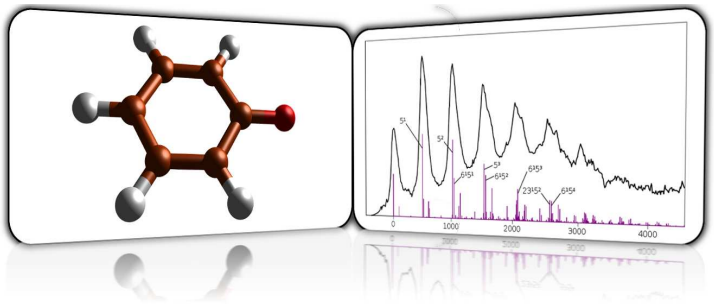
Università degli studi di Napoli
Federico II

FACOLTÀ DI SCIENZE MATEMATICHE, FISICHE E NATURALI

Dottorato di Ricerca in Scienze Chimiche
XXI ciclo

Development and Application of time dependent and time independent models for the study of spectroscopic properties in compounds of biological interest



Julien Bloino

Tutore
Prof. Vincenzo Barone

Relatore
Prof. Filomena Sica

Coordinatore
Prof. Aldo Vitagliano

Novembre 2008

Abstract

In this thesis, we propose a general and effective approach to compute vibrationally-resolved electronic spectra from first principles. This method is integrated in a versatile quantum chemical computational package and offers a complete "in silico" procedure starting from the geometry optimization to the generation of the spectrum.

The theoretical background and methods to evaluate the overlap integrals are presented, along with a discussion of strategies for an efficient evaluation of spectra of large systems, which features a huge number of possible vibronic transitions. The presented procedure relies on the general-purpose method to select *a priori* the transitions that should be calculated by estimating their probability. The implemented method uses a partition of the transitions by groups called *classes*, which permits the usage of several computational schemes to speed up the calculations. The details of the procedure and the possibilities of fine-tuning of the calculations are presented, as well as an insight into its internal workout.

The integrated approach to compute vibrationally resolved optical spectra can be applied to a large variety of systems ranging from small molecules in the gas phase to macrosystems in condensed phases, whenever nonadiabatic couplings are negligible and the harmonic approximation is reliable. The given examples of absorption spectrum of $S_1 \leftarrow S_0$ electronic transitions of anisole, photodetachment spectrum of $SF_6{}^-$, emission $T_1 \rightarrow S_0$ phosphorescence spectrum of chlorophyll *c2*, UV spectrum of acrolein in the gas phase and aqueous solution, a photoelectron spectrum of adenine adsorbed on the Si(100) surface, and porphyrin are chosen to illustrate the possibilities of the procedure and some of its characteristics.

It is shown that despite the fact that our computational scheme has been tailored for large systems, it can be utilized as well to generate high quality spectra for small systems. Moreover, good quality spectra can be effectively computed even for large systems with hundreds of normal modes, paving the route to spectroscopic studies of systems of direct biological and/or technological interest.

Acknowledgement

I am very thankful to my supervisor, professor Vincenzo Barone, who welcomed me in his group. The three years of this thesis have been really enriching for me. I have really appreciated his availability and help which allowed me to smoothly carry out this project.

I also am very thankful to Dr. Małgorzata Biczysko. Her knowledge and our fruitful discussions have been really important for me to improve my understanding of spectroscopic phenomena. Her tests and advice have permitted to enhanced the stability and the possibilities offered by the procedure developed during this work. I also thank her for her contribution to the mode-specific scaling scheme and the study of anisole. Finally, I am very grateful for the time she took to correct this document.

I am greatly thankful to Dr. Fabrizio Santoro for his availability and his help in the development of the procedure. I have really appreciated our discussions and I am very grateful for the time he took to help me correct and improve the quality of the procedure. I thank him for his contribution to the study of porphyrin and his advice for the writing of the present manuscript.

I also want to express my gratitude to professor Orlando Crescenzi for the interesting exchanges we had on spectroscopy techniques and for having helped me to clarify some concepts presented in this document.

I am thankful to Dr. Giovanni Morelli, Michele Pavone and Nadia Rega for helping me to understand the working of GAUSSIAN and efficiently develop the procedure inside it.

I wish to thank professor Filomena Sica for our discussions during the writing of this document.

I am very grateful to the members of the Laboratory of Structure and Dynamics of Molecules for their availability and kindness which permitted me to quickly become well integrated in their group.

Finally, I wish to thank my family and friends for their support and understanding during the writing of the mansuscript.

Contents

Introduction **11**
The theoretical spectrum generation as an important tool for practical cases 11
A brief overview of the available methods to compute absorption and emission UV-visible spectra 11
Aim of this work 11

1 A theoretical approach to compute UV-visible spectra **15**
1.1 Field of application of the theoretical method 15
1.2 Analysis of the transition probability 19
1.3 The transition dipole moment integral 20
1.4 A particular case of the Franck-Condon integrals : $\langle \mathbf{0'} \mid \mathbf{0''} \rangle$ 27

2 Strategies to compute the Franck-Condon integral $\langle \boldsymbol{v'} \mid \boldsymbol{v''} \rangle$ **31**
2.1 Introduction 31
2.2 Analytic formulae and recursive methods 34
2.3 The general analytic approach of Sharp and Rosenstock 36
2.4 A recursive method based on the Sharp-Rosenstock approach 39
2.5 Aiming at an efficient calculation 43
2.6 Indexing and recovering overlap integrals 44
2.7 Prescreening of the overlap integrals 50
2.8 The evaluation method *FCClasses* 56
2.8.1 Extending the evaluation method to the Herzberg-Teller calculations 60
2.9 Evaluating the reliability of the prescreening: the sum rules 62

3 Implementation inside GAUSSIAN **67**
3.1 Introduction 67
3.2 Overview of GAUSSIAN and the interaction between FRANCK and the rest of the software 69
3.3 Extraction of the required data and adjustments 74
3.3.1 A simple scheme for anharmonicity 77
3.4 Control of the output 79
3.5 Calculation of the required matrices 84
3.6 General structure to organize the computation of the transition dipole moment integrals 87
3.7 Management of the classes and calculation of the transition dipole moment integrals 90
3.8 Managing the storage and the indexing of the overlap integrals 95
3.9 Conclusion and discussion 98

4 Examples of application of the procedure **103**
4.1 Introduction 103
4.2 Phenoxyl 104
4.2.1 Introduction 104
4.2.2 Computational details 104
4.2.3 Discussion 106

4.3 Sulfur hexafluoride . . . 112
4.3.1 Introduction . . . 112
4.3.2 Computational details . . . 112
4.3.3 Discussion . . . 112
4.4 Chlorophyll *c2* . . . 115
4.4.1 Introduction . . . 115
4.4.2 Computational details . . . 115
4.4.3 Discussion . . . 115
4.5 Acrolein . . . 120
4.5.1 Introduction . . . 120
4.5.2 Computational details . . . 120
4.5.3 Discussion . . . 120
4.6 Adenine on silicon (100) . . . 124
4.6.1 Introduction . . . 124
4.6.2 Computational Details . . . 124
4.6.3 Discussion . . . 124
4.7 Porphyrin . . . 127
4.7.1 Introduction . . . 127
4.7.2 Computational details . . . 128
4.7.3 Results and discussion . . . 129
4.8 Anisole . . . 136
4.8.1 Introduction . . . 136
4.8.2 Quantum mechanical calculations . . . 137
4.8.3 Results and Discussion . . . 137

Conclusion **163**

A General formula of the transition dipole moment integral **165**

B Details for the calculation of the overlap integral $\langle\, \mathbf{0}' \mid \mathbf{0}'' \,\rangle$ **167**

C Calculation of the power series $f(\mathbf{T}, \mathbf{U})$ **169**

D Calculation of the transition dipole moment integral in the case of D2 **171**

E List of the keywords currently recognized by FRANCK **177**

F Supplementary data for the study of the $^2B_1 \rightarrow {}^2A_2$ transition of the phenoxyl radical **181**
F.1 GAUSSIAN Input files . . . 181
F.2 Reoriented/superposed structures . . . 182

Conventions used in this document

In this document several mathematical symbols will be used to represent the physical and chemical concepts needed to grasp the underlying theory behind the computation of vibronically resolved UV-visible spectra. Their meaning will be explained as they are introduced. However, to ease the reading of this document, this chapter intends to briefly sum up the conventions used throughout the document.

Typographic conventions	
$\mathbf{A}$	represents a matrix or vector. The character is upper-case.
$\underline{\mathbf{A}}$	represents a submatrix or subvector of $\mathbf{A}$.
A_i	represents an element i of a vector $\mathbf{A}$. Occasionally, the notation $\mathrm{A}(i)$ can be preferred for clarity.
A_{ij}	represents an element i, j of a matrix $\mathbf{A}$. Occasionally, the notation $\mathrm{A}(i, j)$ can be preferred for clarity.
$\mathbf{A}^{\mathrm{T}}$	represents the transpose of matrix/vector $\mathbf{A}$.
$'$	placed next to a variable/symbol, the latter describes the initial state.
$''$	placed next to a variable/symbol, the latter describes the final state.
Main mathematical symbols	
N_A	represents the number of atoms of the molecule.
N	represents the number of normal modes in the molecule. For linear molecules, $N = 3N_A - 5$ while for the others, $N = 3N_A - 6$.
$\mathbf{Q}$	is a vector representing the normal coordinates of one electronic state.
$\mathbf{q}$	is a vector representing the dimensionless normal coordinates of one electronic state.
Ψ	is the molecular wave function.
ψ_e	is the electronic wave function.
ψ_n	is the nuclear wave function.
ψ_v	is the vibrational wave function.
ψ_{v_i}	is the monodimensional vibrational wave function associated to the normal mode i.
$\boldsymbol{v}$	is a vector of quantum numbers representing a vibrational state. It is used as an equivalent to ψ_v in matrix notation in harmonic approximation.

Introduction

The theoretical spectrum generation as an important tool for practical cases

For years there has been a constant interplay between molecular spectroscopy and computational chemistry. Computed data have become crucial for the interpretation of experimental results and, conversely, accurate spectroscopic measurements have been used as benchmarks to validate theoretical approaches. For small molecular systems, state-of-the-art quantum mechanical (QM) approaches to computational spectroscopy are at present capable of providing results comparable to the most accurate experimental measurements [1]. Among the most challenging examples, we can mention theoretical spectroscopic studies of systems with several interacting electronic states, based on the evaluation of accurate ab initio post-Hartree-Fock potential energy surfaces (PES), and variational calculations of rovibronic energy levels beyond the Born-Oppenheimer approximation [2, 3]. The high accuracy of the results achievable for small systems clearly demonstrates the potentiality of computational chemistry experiments to become key tools for the prediction and understanding of spectroscopic properties of all kinds of molecular systems: this reflects a more general trend, since nowadays the experimental characterization of new systems relies more and more on computational approaches, e.g. for the evaluation and rationalization of structural, energetic, electronic and dynamic features [4–6]. Nevertheless, up to very recently direct comparisons between experimental and computed spectroscopic data have been rather scarce. As a matter of fact, even for small molecular systems comparisons have mostly been restricted to energy levels (vibrational, electronic, rovibrational), which as a rule must be extracted from the experimental data by a non trivial interpretation (involving at the very least band assignment). Alternatively, model Hamiltonians have been extensively applied to extract experimental spectroscopic constants and to simulate spectra using trial constants - an approach which in a sense is equally far removed from experimental spectroscopy, as from electronic structure methods. However, the need for integrated approaches, capable of accurately simulating optical spectra, but at the same time easily accessible to non-specialists, is strongly felt. Such tools would allow for the exploitation of the recent and ongoing developments that are taking place in the field, resulting in easy and, ideally, automatic vis-à-vis comparison between experimental and theoretical results.

A brief overview of the available methods to compute absorption and emission UV-visible spectra

Many attempts have been made toward a fully ab initio simulation of vibronic spectra of medium to large molecular systems, by both time-dependent and time-independent approaches [7,8]. Eigenstate-free time-dependent methods are the main (when not the only) route to deal with systems for which eigenstate calculations are unfeasible, as is the case for strongly non-adiabatic systems involving conical intersections [7], or for systems propagating on highly anharmonic potential energy surfaces (PES) [9,10]. On the other hand, for large and semi-rigid molecules, when non-adiabatic couplings are negligible and the harmonic approximation is reliable, vibronic eigenstates are known once a harmonic analysis of the relevant PES is performed (a step required for the characterization of the PES, independently of the method adopted to compute the spectrum). Within the simplest zero-order harmonic approximation it is assumed that the PES of the initial and final state do not differ in shape, so that the vibrational levels are identical. Obviously such an approach can only provide a very rough estimation of the real spectrum, and in general most vibrational transitions are actually missed. A significant improvement is represented by the Linear Coupling Method (LCM) [11]. Here the multidimensional coupling constants are obtained from the ground state frequencies and normal modes, and the excited state energy gradients evaluated at the ground state geometry. Thus, the approach does not require computation of the excited state equilibrium geometry, frequencies and normal modes, which have only recently become feasible for large systems. However, the Linear Coupling Method does not take into account changes in vibrational frequencies and/or in normal modes between the excited and ground state. More accurate approaches require a detailed knowledge of the multidimensional potential energy surfaces (PES) of both electronic states or, within the harmonic approximation, at least computation of equilibrium geometry structures and vibrational properties. Till recently, computations of vibronic spectra have been limited to small systems or approximated approaches, mainly a consequence of the difficulties in obtaining accurate descriptions of excited electronic states of polyatomic molecules. Recent developments in electronic structure theory for excited states [12] within the Time-Dependent Density Functional Theory(TD-DFT) [13,14] and resolution-of-the-identity approximation of Coupled Cluster theory(RI-CC2) [15] have paved the route toward the simulation of spectra for significantly larger systems.

Aim of this work

In this work, we aimed at developing a fast and general-purpose procedure to compute vibronically-resolved ultraviolet-visible spectra. To fulfill our objective, the principle stated by Franck [16] and formalized by Condon [17] was well adapted. However, the Franck-Condon principle is not sufficient to permit the generation of an electronic spectrum. As a matter of fact, we first had to develop a thorough theoretical method that could be efficient and versatile. Once it has been sufficiently consistent, its encoding could start. This part cannot be neglected because it plays also an important role in the overall speed of the calculation program. As a result, we will also give a substantial place to the programming strategies to fully take advantage of modern computers performances and not diminish the intrinsic qualities of our method. In addition to being a concrete application of a theoretical method, the program must allow the users to customize the calculations

to their needs. In this project, we have chosen to build an easy-to-use tool that could be handled by non-specialists without requiring an in-depth understanding of the underlying principles it relies on. A straightforward approach would be to create a stand-alone software but this has several disadvantages. The main one is that such a program is not self-contained and must be able to interface it with existing calculation packages to extract the data it needs. Consequently, it is easier and more efficient to integrate directly the procedure in a quantum mechanical computational package. This gives access to the internal functionalities of these softwares as well as the quantities they have computed. In this work, we chose to implement our method inside the well-known computational package GAUSSIAN [18].

Once the procedure has become stable enough, we applied it to several test cases. From these experiments, we were able to refine our model, from a theoretical perspective [19] as well as from a computational perspective to improve its efficiency. Additionally, these models help us to improve the range of application of the program as well as the possibilities to fine-tune its results. Finally, we used it for some specific studies [19–21].

The document is divided in three parts organized in four chapters. In the first chapter, we will discuss the main approximations underlying our work and will present the exact calculation of the overlap integral between the vibrational ground states of two electronic states. In the second chapter, we will focus on the general calculations of the overlap integrals between arbitrary vibrational states. The strategies of computations will be discussed from two perspectives. In a first time, we will present the theoretical method we adopted in our case to generate efficiently the vibronic spectra. Then, we will consider the implications of such an approach from a programming point of view. This will pave the way to our third chapter which will describe the procedure inside GAUSSIAN. The presentation will be done with an emphasis on its usage and the possibilities it offers to parametrize the calculation of the spectra and the output. When relevant, we will explain the programming strategies we adopted to maximize the speed of the procedure. Finally, in the last chapter, we will present some applications of our procedure or method with different systems and conditions. These examples will highlight the influence of some parameters on the spectrum generation (sulfur hexafluoride anion, chlorophyll *c2*) and the possibilities offered by the full integration in a quantum chemical computational package such as GAUSSIAN (acrolein, adenine on a silicon surface). We will also present a full study of the UV-vis spectra of two systems with respect to experimental results: porphyrin and anisole.

Chapter 1

A theoretical approach to compute UV-visible spectra

1.1 Field of application of the theoretical method

In this document, all discussions will be done assuming a certain number of approximations. The first and most important one is the Born-Oppenheimer (BO) approximation, related to the concept of separability of the potential energy surface of each electronic state. In our approach, we will always assume that it does not break down. Because the nuclei are far heavier than the electrons, they move more slowly. Hence, it can be assumed, with a good approximation, that the electrons are interacting and moving in a field of fixed nuclei. A direct consequence of this approximation is that the total Hamiltonian $\mathcal{H}$ of the Schrödinger equation can be written as the sum of an electronic Hamiltonian $\mathcal{H}_e$ and a nuclear Hamiltonian $\mathcal{H}_n$ where the swift electrons are represented by an average field in which the nuclei move. Another consequence of this approximation, which will be widely used further in this document is the possibility to rewrite the total molecular wave function Ψ as the product of an electronic wave function ψ_e and a nuclear wave function ψ_n. The simulated spectra, in our case, are generated from vibrational transitions so the electronic parts of the wave function and the Hamiltonian will not be explicitly treated and only the nuclear Shrödinger equation needs to be considered.

The BO approximation is however not sufficient as the nuclear wave function describes complex movements involving vibrations, rotations and translations. Since we focus on vibrational displacements, it is necessary to find a way to separate, as far as possible, the vibrations from the other sources of nuclear motions. It is possible to do so using the Sayvetz conditions [22], also known as the Eckart conditions [23]. These conditions are physically less intuitive than the Born-Oppenheimer approximation and also imply some restrictions on the relevant systems. To define them, let us formulate the nuclear Hamiltonian in a classical approach. In the case of an isolated molecule, the nuclear Hamiltonian $\mathcal{H}_n$ is correlated to the kinetic energy $\mathcal{V}$ which can be expressed as

$$\mathcal{V} = \sum_{a=1}^{N_A} \frac{1}{2} m_a \mathbf{V}_a \cdot \mathbf{V}_a \tag{1.1}$$

where N_A is the number of atoms in the molecule and m_a and V_a are the mass and the velocity

of nucleus a, respectively.

At the center of mass of the molecule, let us define a set of orthonormal coordinates from which the position of each nucleus is measured. This rotating frame of reference $\mathcal{R}_l$ is at a distance $\mathbf{R}$ from the origin of the fixed spatial frame of reference $\mathcal{R}$ and has an angular velocity $\boldsymbol{\omega}_r$. From a physical point of view, $\mathbf{R}$ represents the translation of the molecule in the cartesian space while $\boldsymbol{\omega}_r$ represents its rotation.

In the frame of reference $\mathcal{R}_l$, the position of some nucleus a is defined by the vector $\mathbf{r}_a$ and its displacement is given by $\boldsymbol{\delta}_a = \mathbf{r}_a - \mathbf{r}_{a_0}$ where $\mathbf{r}_{a_0}$ is the equilibrium position. The set of instantaneous displacements $\boldsymbol{\delta}_a$ corresponds to the vibrations of the atoms in the molecule.

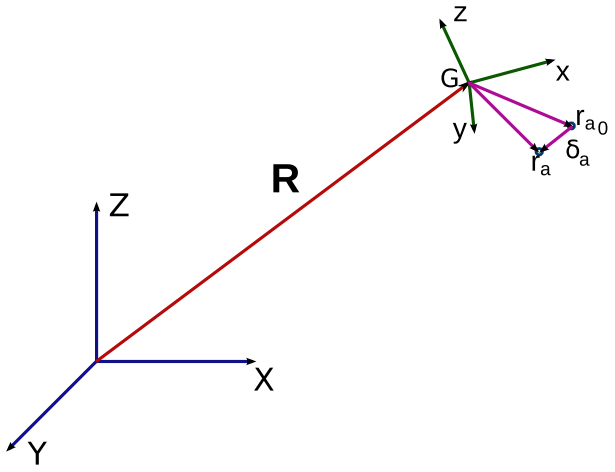


Figure 1.1: Diagram of the two frames of reference with a nucleus a. The fixed spatial frame of reference $\mathcal{R}$ is defined by the set (X,Y,Z) and the rotating frame of reference $\mathcal{R}_l$ is defined by (x,y,z).

The velocity of the nucleus a in the spatial frame of reference $\mathcal{R}$ is:

$$\mathbf{V}_a = \frac{d\mathbf{R}}{dt} + \boldsymbol{\omega}_r \times \mathbf{r}_a + \frac{d\boldsymbol{\delta}_a}{dt} \tag{1.2}$$

Using equation 1.2, the kinetic energy can be written

$$\begin{aligned}\mathcal{V} = \frac{1}{2}\sum_{a=1}^{N_A}\Bigg\{ & m_a\left(\frac{d\mathbf{R}}{dt}\right)^2 + m_a\big(\boldsymbol{\omega}_r \times \mathbf{r}_a\big)^2 + m_a\left(\frac{d\boldsymbol{\delta}_a}{dt}\right)^2 + 2m_a\frac{d\mathbf{R}}{dt}\cdot\big(\boldsymbol{\omega}_r \times \mathbf{r}_a\big) \\ & + 2m_a\frac{d\mathbf{R}}{dt}\cdot\frac{d\boldsymbol{\delta}_a}{dt} + 2m_a\big(\boldsymbol{\omega}_r \times \mathbf{r}_a\big)\cdot\frac{d\boldsymbol{\delta}_a}{dt}\Bigg\}\end{aligned} \tag{1.3}$$

The first three terms of the right-hand side of equation 1.3 represent the translational, rotational and vibrational movements, respectively. The last three terms correspond to the interactions between these movements.

To be able to separate the vibrational wave function from the rotational and translational ones in the nuclear wave function, it is necessary to cancel or at least highly reduce the value of these last terms.

As a consequence of equation 1.3, the nuclear Hamiltonian is defined with $3N_A + 6$ coordinates, one for each Cartesian coordinate of each nucleus ($3 \times N_A$ coordinates), 3 for the translation vector $\mathbf{r}_a$ and 3 for the rotation vector $\boldsymbol{\omega}_r$ each. However, the molecule has $3N_A$ degrees of freedom so it is possible to set 6 conditions among the coordinates. These correspond to the Eckart conditions.

The first condition, which corresponds to 3 conditions in Cartesian coordinates, is

$$\sum_{a=1}^{N_A} m_a \mathbf{r}_a = \mathbf{0} \tag{1.4}$$

and implies that *the center of mass does not shift during a molecular vibration.*

By differentiating this relation with respect to time, another consequence of this statement can be expressed.

$$\frac{d}{dt}\left(\sum_{a=1}^{N_A} m_a \mathbf{r}_a\right) = \sum_{a=1}^{N_A} m_a \left(\boldsymbol{\omega}_r \times \mathbf{r}_a + \frac{d\boldsymbol{\delta}_a}{dt}\right) = \boldsymbol{\omega}_r \times \sum_{a=1}^{N_A} m_a \mathbf{r}_a + \sum_{a=1}^{N_A} m_a \frac{d\boldsymbol{\delta}_a}{dt} = \mathbf{0} \tag{1.5}$$

$$\Rightarrow \sum_{a=1}^{N_A} m_a \frac{d\boldsymbol{\delta}_a}{dt} = \mathbf{0} \tag{1.6}$$

Hence, the first Eckart condition imposes that there is *no linear momentum of vibration.*

From this expression, both interactions between the translation and the rotation, and between the translation and the vibrations are null. Only the sixth term of equation 1.3, corresponding to the interaction between rotation and vibrations remain. This term is rather complex and cannot be easily cancelled out, even with the last three remaining conditions to set.

The second Eckart condition states that *no angular momentum is created during a molecular vibration.* Hence, any global rotation induced by a vibration is counterbalanced by the rotating system. This can be formulated as

$$\sum_{a=1}^{N_A} m_a \mathbf{r}_{a_0} \times \mathbf{r}_a = \mathbf{0} \tag{1.7}$$

To link this equation to the term of rovibrational interaction, the latter must be slightly rewritten. Using the scalar triple product, a permutation of the vectors is possible.

$$\sum_{a=1}^{N_A} m_a \big(\boldsymbol{\omega}_r \times \mathbf{r}_a\big) \cdot \frac{d\boldsymbol{\delta}_a}{dt} = \sum_{a=1}^{N_A} m_a \Big(\mathbf{r}_a \times \frac{d\boldsymbol{\delta}_a}{dt}\Big) \cdot \boldsymbol{\omega}_r = \boldsymbol{\omega}_r \cdot \sum_{a=1}^{N_A} m_a \big(\mathbf{r}_{a_0} + \boldsymbol{\delta}_a\big) \cdot \frac{d\boldsymbol{\delta}_a}{dt} \tag{1.8}$$

Derivating the second condition of Eckart with respect to time, one obtains

$$\begin{aligned} \frac{d}{dt}\sum_{a=1}^{N_A} m_a \mathbf{r}_{a_0} \times \mathbf{r}_a &= \sum_{a=1}^{N_A} m_a \left[\frac{d\mathbf{r}_{a_0}}{dt} \times \mathbf{r}_a + \mathbf{r}_{a_0} \times \frac{d\mathbf{r}_a}{dt}\right] \\ &= \sum_{a=1}^{N_A} m_a \left[(\boldsymbol{\omega}_r \times \mathbf{r}_{a_0}) \times \mathbf{r}_a + \Big(\mathbf{r}_{a_0} \times (\boldsymbol{\omega}_r \times \mathbf{r}_a) + \mathbf{r}_{a_0} \times \frac{d\boldsymbol{\delta}_a}{dt}\Big)\right] = \mathbf{0} \end{aligned} \tag{1.9}$$

Using the anticommutative properties of the cross product, the first term of the right-hand side of equation 1.9 becomes

$$(\boldsymbol{\omega}_r \times \mathbf{r}_{a_0}) \times \mathbf{r}_a = -\mathbf{r}_a \times (\boldsymbol{\omega}_r \times \mathbf{r}_{a_0}) = \mathbf{r}_a \times (\mathbf{r}_{a_0} \times \boldsymbol{\omega}_r) \tag{1.10}$$

The cross product satisfying the Jacobi identity[a], the first two terms of the right-hand side of equation 1.9 can be replaced by a single cross product given by the relation

$$\boldsymbol{\omega}_r \times (\mathbf{r}_a \times \mathbf{r}_{a_0}) = -\big[\mathbf{r}_{a_0} \times (\boldsymbol{\omega}_r \times \mathbf{r}_a) + \mathbf{r}_a \times (\mathbf{r}_{a_0} \times \boldsymbol{\omega}_r)\big] \tag{1.11}$$

Applying the second Eckart condition, equation 1.9 is reduced to

$$\sum_{a=1}^{N_A} m_a \mathbf{r}_{a_0} \times \frac{d\boldsymbol{\delta}_a}{dt} = \mathbf{0} \tag{1.12}$$

It is now possible to simplify equation 1.8 as

$$\sum_{a=1}^{N_A} m_a \big(\boldsymbol{\omega}_r \times \mathbf{r}_a\big) \cdot \frac{d\boldsymbol{\delta}_a}{dt} = \boldsymbol{\omega}_r \sum_{a=1}^{N_A} m_a \boldsymbol{\delta}_a \cdot \frac{d\boldsymbol{\delta}_a}{dt} \tag{1.13}$$

Unfortunately, it is not possible to completely remove the rovibrational interaction. However, this term, called Coriolis energy, is very small and is generally included in the rotational movement of the molecule. Consequently, if the Eckart conditions are met, it is possible to write the nuclear wave function as

$$\psi_n = \psi_r \psi_v \tag{1.14}$$

where ψ_r is the rotational wave function and ψ_v the vibrational wave function. As the translational motion is completely separated from the rotational and vibrational ones, it can be treated separately and will be ignored in our treatment of the radiative transitions.

Finally, a last important approximation applied throughout this work is the harmonic oscillator approximation. The potential surface near a local minimum can be modelled by a set of simple harmonic oscillators corresponding to the vibrational modes of the molecule. In this setting, each molecular vibration can be described as a linear combination of independent coordinates ($Q_1, \ldots, Q_N$, $N = 3N_A - 6$[b]) called normal coordinates, that will be detailed further. It is then possible to express the multidimensional vibrational wave function ψ_v as a product of N monodimensional functions $\psi_{v_i}(Q_i)$.

It should be noted that the approach presented here can be extended to deal with anharmonic

[a] This identity can be shown using the vector triple product identity:

$$\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = \mathbf{b}(\mathbf{a} \cdot \mathbf{c}) - \mathbf{c}(\mathbf{a} \cdot \mathbf{b})$$

Such a relation can also be applied to $\mathbf{b} \times (\mathbf{c} \times \mathbf{a})$ and $\mathbf{c} \times (\mathbf{a} \times \mathbf{b})$:

$$\mathbf{b} \times (\mathbf{c} \times \mathbf{a}) = \mathbf{c}(\mathbf{b} \cdot \mathbf{a} - \mathbf{a}(\mathbf{b} \cdot \mathbf{c}))$$

$$\mathbf{c} \times (\mathbf{a} \times \mathbf{b}) = \mathbf{a}(\mathbf{c} \cdot \mathbf{b} - \mathbf{b}(\mathbf{c} \cdot \mathbf{a}))$$

Since the dot product is communative, the following equalities can be written:

$$\mathbf{a} \cdot \mathbf{b} = \mathbf{b} \cdot \mathbf{a} \quad ; \quad \mathbf{a} \cdot \mathbf{c} = \mathbf{c} \cdot \mathbf{a} \quad ; \quad \mathbf{b} \cdot \mathbf{c} = \mathbf{c} \cdot \mathbf{b}$$

Using this property, it is straightforward to obtain the Jacobi identity:

$$\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) + \mathbf{b} \times (\mathbf{c} \times \mathbf{a}) + \mathbf{c} \times (\mathbf{a} \times \mathbf{b}) = \mathbf{0}$$

[b] $N = 3N_A - 5$ for linear molecules

oscillators, making necessary corrections to adapt the formulae to this model. However, this treatment goes beyond the scope of this document and will not be discussed here.

To make the formulae used later in this document as clear as possible, the convenient Dirac notation will be used.

1.2 Analysis of the transition probability

In this section, we will present the formulae used to compute the absorption and emission spectra of UV and visible radiations. While the analysis of the transitions is necessary to generate these spectra, a complete description of the light matter interaction is not indispensable to the understanding of the present document. Such an interaction can be treated with two different approaches. The simpler one is called *semiclassical* and uses a classical electromagnetic field to describe the radiation and quantum mechanics for the molecule. The second one, the *quantum field* theory, treats both field and system with quantum mechanics. Whatever the choice of the method to describe the interaction, the time-dependent perturbation is necessary. In this discussion, however, we will restrict our study to some important concepts, focusing on the case of absorption. More thorough and detailed analyses can be found in books about spectroscopy such as "Vibrational States" by S. Califano [24] or "Molecular Spectra and Molecular Structure" by G. Herzberg. [25].

When a system is thermically excited, its vibrational state before a radiative transition can be different from the ground state (case of hotbands). Hence, the calculation of an absorption (or emission) spectrum will require the knowledge of the transition probabilities between each possible initial and final states.

As a brief description, the transition probability from one energy eigenstate to a distribution of eigenstates is assumed to form, in general, a continuum. The measure of the transition probability is supposed to be done after any possible transition has occured. In this setup, the transition rate, which is the transition probability per time unit, between an initial state i and a final state f is given by the Fermi's golden rule [26, 27]:

$$R_{if} = \frac{2\pi}{\hbar}|\mathrm{V}_{if}|^2\rho(E_f) \tag{1.15}$$

where $\rho(E_f)$ is the density of states which describes the distribution of final states and V_{if} is the matrix element connecting the states.

V_{if} can also be written $\langle\ \Psi_i\ |\ \mathcal{H}_{int}\ |\ \Psi_f\ \rangle$ where Ψ_i and Ψ_f are the wave functions of the initial state and final state, respectively, and $\mathcal{H}_{int}$ is the interaction Hamiltonian. In the case of UV-visible radiations, $\mathcal{H}_{int}$ represents the interaction between the molecule and an external electromagnetic field. The Hamiltonian will then contain multiple moments of the molecule and powers of the magnetic and electric field. The three most relevant moments are the electric dipole moment, the electric quadrupole moment and the magnetic dipole moment. However, since the electric quadrupole and magnetic dipole moments are very small with respect to the first one, only the electric dipole moment is generally considered in interactions between light and matter. As a

consequence, $\mathcal{H}_{int}$ can be replaced with the relation

$$\mathcal{H}_{int} = (\overrightarrow{\boldsymbol{\mu}_E} \cdot \vec{e})\, E_0 \tag{1.16}$$

where $\overrightarrow{\boldsymbol{\mu}_E}$ is the electric dipole moment and $\vec{e}$ is a dimensionless unit vector indicating the direction of the electric field whose intensity is E_0

Hence, equation 1.15 can be written

$$R_{if} = \frac{2\pi E_0^2}{\hbar} \left| \langle\, \Psi_i \mid \overrightarrow{\boldsymbol{\mu}_E} \cdot \vec{e} \mid \Psi_f \,\rangle \right|^2 \rho(E_f) \tag{1.17}$$

The transition between the vibronic states i and f occurs when the molecule absorbs a photon of energy $\hbar\omega$ corresponding to the difference of energy between them. It should be noted that the probability of a transition from the state i to the state f is sharply centered on $E_f = E_i + \hbar\omega$. As a consequence, the density of state can be written $\rho(E_f = E_i + \hbar\omega)$ and is more commonly replace by a Dirac function $\delta(E_f - E_i - \hbar\omega)$.

We define the absorption cross section as the rate of photon absorption per molecule and per unit of radiant energy as:

$$\sigma(\omega) = \frac{4\pi^2\omega}{c} \left| \langle\, \Psi_i \mid \overrightarrow{\boldsymbol{\mu}_E} \cdot \vec{e} \mid \Psi_f \,\rangle \right|^2 \delta(E_f - E_i - \hbar\omega) \tag{1.18}$$

The stick spectrum from a given vibronic state i can be calculated as a sum of the absorption coefficient $\sigma(\omega)$ over all possible final vibronic state f. If temperature is taken into account, the different initial states and their Boltzmann population ρ_i are taken into account in equation 1.18:

$$\sigma(\omega) = \frac{4\pi^2\omega}{c} \sum_i \rho_i \sum_f \left| \langle\, \Psi_i \mid \overrightarrow{\boldsymbol{\mu}_E} \cdot \vec{e} \mid \Psi_f \,\rangle \right|^2 \delta(E_f - E_i - \hbar\omega) \tag{1.19}$$

where $\sum_f$ represents all possible final states for a given energy of transition.

If we assume that the orientation of the studied system is completely random, then the Cartesian components of the electric dipole moment are the same, so that we can average them, writing:

$$\left|\boldsymbol{\mu}\right|^2 = 3\left|\boldsymbol{\mu}_E(x)\right|^2$$

$$\sigma(\omega) = \frac{4\pi^2\omega}{3c} \sum_i \rho_i \sum_f \left| \langle\, \Psi_i \mid \boldsymbol{\mu} \mid \Psi_f \,\rangle \right|^2 \delta(E_f - E_i - \hbar\omega) \tag{1.20}$$

For the emission spectra, we will use the following formula:

$$\sigma(\omega) = \frac{4\omega^3}{3\hbar c^3} \sum_i \rho_i \sum_f \left| \langle\, \Psi_i \mid \boldsymbol{\mu} \mid \Psi_f \,\rangle \right|^2 \delta(E_f - E_i + \hbar\omega) \tag{1.21}$$

1.3 The transition dipole moment integral

The absorption or emission spectrum can be calculated using the relations given in equations 1.20 and 1.21, respectively. From the equations of the stick spectrum, it appears that the probability of a transition is proportional to the square of the transition dipole moment integral $\langle\, \Psi_i \mid \overrightarrow{\boldsymbol{\mu}_E} \cdot \vec{e} \mid \Psi_f \,\rangle$,

and the intensity of a spectrum line depends mainly on the transition probabilities for a given incident energy. Evaluating correctly each probability of transition is an essential prerequisite to the correct computation of the complete spectrum.

To simplify the equations, some conventions will be used. The initial state i will be referred with the single quote ($'$) and the final state with the double quote ($''$). For convenience, this notation is slightly different from the spectroscopic conventions as the relative energies of the initial and final states are not taken into account in our notation.

In this section the Franck-Condon principle will be used to evaluate the probability of a transition with respect to the overlap of the vibrational wave functions of the initial and final vibrational states. To summarize it, Franck [16] firstly proposed that during a transition, the electron jump in a molecule takes place in such a short time that the relative positions and velocities of the nuclei are nearly unaltered by the molecular vibrations. Following this principle, the most intense transition will be from the chosen initial state to a final state being at a minimum of the lowest potential surface vertically upward as shown in figure 1.2. The implications of this approximation and its domain of application will be discussed below.

However, before using it, it is necessary to simplify the transition dipole moment integral using simple schemes first. Let us separate the electric dipole moment $\boldsymbol{\mu}$ into an electronic part $\boldsymbol{\mu}_e$ and a nuclear part $\boldsymbol{\mu}_n$ depending on the electrons and the nuclei of the system, respectively. From the Born-Oppenheimer approximation, the molecular wave function Ψ can be written as a product of an electronic wave function ψ_e and a nuclear wave function ψ_n. Consequently, the transition dipole moment can be developed into two integrals:

$$\langle \Psi' \mid \boldsymbol{\mu} \mid \Psi'' \rangle = \langle \psi'_e\psi'_n \mid \boldsymbol{\mu}_e \mid \psi''_e\psi''_n \rangle + \langle \psi'_e\psi'_n \mid \boldsymbol{\mu}_n \mid \psi''_e\psi''_n \rangle \tag{1.22}$$

Since $\boldsymbol{\mu}_n$ is independent from the coordinates of the electrons, the second term of equation 1.22 can be written

$$\langle \psi'_e\psi'_n \mid \boldsymbol{\mu}_n \mid \psi''_e\psi''_n \rangle = \langle \psi'_n \mid \boldsymbol{\mu}_n \mid \psi''_n \rangle \langle \psi'_e \mid \psi''_e \rangle \tag{1.23}$$

The electronic wave functions of different electronic states are orthogonal to each other so $\langle \psi'_e \mid \psi''_e \rangle = 0$. Equation 1.22 is now reduced to the first integral:

$$\langle \Psi' \mid \boldsymbol{\mu} \mid \Psi'' \rangle = \langle \psi'_e\psi'_n \mid \boldsymbol{\mu}_e \mid \psi''_e\psi''_n \rangle \tag{1.24}$$

The integral $\langle \psi'_e \mid \boldsymbol{\mu}_e \mid \psi''_e \rangle$ corresponds to the electronic transition dipole moment and will be written $\boldsymbol{\mu}_{if}$ from now on. As a consequence, the transition dipole moment depends now on the nuclear wave functions and the electronic transition moment:

$$\langle \Psi' \mid \boldsymbol{\mu} \mid \Psi'' \rangle = \langle \psi'_n \mid \boldsymbol{\mu}_{if} \mid \psi''_n \rangle \tag{1.25}$$

As detailed in section 1.1, we suppose that the Eckart conditions are fulfilled so that the nuclear wave function can be, with a good approximation, separated into a rotation and a vibrational contribution as shown in equation 1.14. In this case, it is possible to neglect the rotation of the molecule (more in-depths and technical details about this approximation can be found in the explanation following eq. [IV,76] in reference [25]). Finally, the transition dipole moment integral

can be written:

$$\langle \Psi' \mid \boldsymbol{\mu} \mid \Psi'' \rangle = \langle \psi'_v \mid \boldsymbol{\mu}_{if} \mid \psi''_v \rangle \tag{1.26}$$

Actually, the relation given in equation 1.26 is not sufficient to calculate the transition moment since there is no general analytic solution for $\boldsymbol{\mu}_{if}$. Previously, the Franck-Condon principle has been described as an approach ignoring the shifts of the nuclei during the transition so that the electronic transition moment is supposed constant. However, we will adopt a slightly more general approximation and take into account a small variation of the positions and velocities of the nuclei during the transition. Physically, it will mean that a broader range of transition can be correctly simulated to improve the quality of the theoretical spectrum.

Since the shifts of the nuclear positions in the molecule are assumed to be fairly low during the transition, it is possible, with a good accuracy, to develop the electronic transition dipole moment in a Taylor series about the equilibrium geometry of the final state. To treat the motions of the nuclei induced by the vibrations, the Cartesian coordinates are not well suited as $3N_A$ coordinates are needed to describe each vibration. We will now introduce a new set of N coordinates $\{Q_i\}$ called normal coordinates. Each normal coordinate Q_i represents the whole of the nuclear motions for a given vibration and so can be expressed as a linear transformation of the Cartesian displacement coordinates $\mathbf{X}$:

$$Q_i = \sum_{k=1}^{N_A} \sqrt{m_k} \sum_{\tau=x,y,z} (\mathrm{L}^+)_{ik\tau} X_{k\tau} \tag{1.27}$$

where m_k is the mass of atom k, $X_{k\tau}$ is one of the Cartesian coordinates of the motion of its nucleus and $(\mathrm{L}^+)_{ik\tau}$ is the coefficient of the linear transformation. In matrix notation, this transformation is simply written as:

$$\mathbf{Q} = \mathbf{L}^+ \mathbf{M}^{1/2} \mathbf{X} \tag{1.28}$$

$\mathbf{L}^+$ is the $(N \times 3N_A)$ matrix pseudoinverse [28] of $\mathbf{L}$ and corresponds to the transformation matrix from mass-weighted Cartesian coordinates to normal coordinates and $\mathbf{M}$ is the diagonal matrix of atomic masses.

The Taylor expansion of the electronic transition moment in the set of normal coordinates is:

$$\boldsymbol{\mu}_{if}(\mathbf{Q}'') \simeq \boldsymbol{\mu}_{if}(\mathbf{Q}''_0) + \sum_{k=1}^{N} \left(\frac{\partial \boldsymbol{\mu}_{if}}{\partial Q''_k} \right)_0 Q''_k + \frac{1}{2} \sum_{k=1}^{N} \sum_{l=1}^{N} \left(\frac{\partial^2 \boldsymbol{\mu}_{if}}{\partial Q''_k \partial Q''_l} \right)_0 Q''_k Q''_L + \dots \tag{1.29}$$

where $\mathbf{Q}''_0$ represents the equilibrium geometry of the final state.

The zeroth order term corresponds to a static electronic transition dipole. This is a direct application of the Franck-Condon principle which was proposed by Franck [16] and mathematically formalized by Condon [17] to calculate the probabilities of vibronic transitions. As a consequence, it is referred as the Franck-Condon approximation. When dealing with fully-allowed transitions ($\boldsymbol{\mu}_{if}(\mathbf{Q}''_0)^2 \gg \mathbf{0}$), this approximation generally gives very good results.

However, in the case of weakly-allowed ($\boldsymbol{\mu}_{if}(\mathbf{Q}''_0) \sim \mathbf{0}$) or dipole-forbidden transitions ($\boldsymbol{\mu}_{if}(\mathbf{Q}''_0) = \mathbf{0}$), this approximation is not describing correctly the overall spectrum, missing the most likely transitions. It is necessary to consider a variation of the dipole moment during the transition. A first approximation is to consider a linear variation of the dipole moment with the normal coordinates.

This corresponds to the first order of the Taylor expansion and is referred as the Herzberg-Teller [29] approximation.

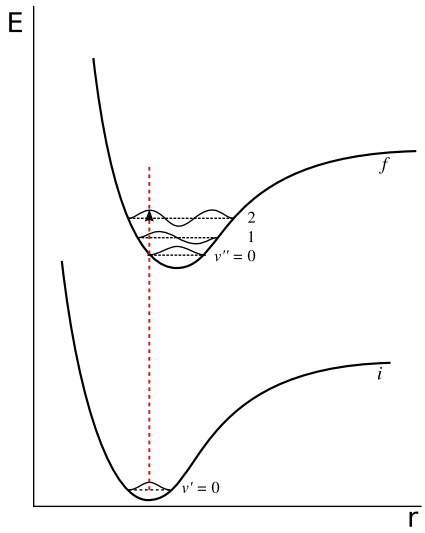


Figure 1.2: The Franck-Condon principle is shown by the vertical dotted line. Following this principle, the highest probability of transition corresponds to $\langle\, \mathbf{0'} \mid \mathbf{2''} \,\rangle$ where the overlap of the wave functions is the highest.

Both approximations are schematically shown in figure 1.2. Taken together, they generally lead to a very good accuracy of the theoretical spectrum of a radiative transition. In some cases [30], the third term of equation 1.29 can be relevant to fully describe the complete spectrum. Because it implies heavy calculations with little benefits in general, it is often not taken into account. In the code developed during this thesis work, we introduced a simple way to take into account parts of this term so this approximation will also be presented here.

The development to a further order of the Taylor expansion of the electronic dipole moment is almost never, to the best of our knowledge, considered in published work, the involved calculations being greatly cumbersome for a very little expected gain. Hence, the electric dipole moment will be set equal to the sum of the first three terms of the right-hand side of equation 1.29 from now on.

Using equation 1.29, the transition dipole moment integral given in equation 1.26 can be written

$$\begin{aligned}\langle\, \Psi' \mid \boldsymbol{\mu} \mid \Psi'' \,\rangle = {} & \boldsymbol{\mu}_{if}(\mathbf{Q}''_0) \,\langle\, \psi'_v \mid \psi''_v \,\rangle \\ & + \sum_{k=1}^{N} \left(\frac{\partial \boldsymbol{\mu}_{if}}{\partial Q''_k} \right)_0 \langle\, \psi'_v \mid Q''_k \mid \psi''_v \,\rangle \\ & + \frac{1}{2} \sum_{k=1}^{N} \sum_{l=1}^{N} \left(\frac{\partial^2 \boldsymbol{\mu}_{if}}{\partial Q''_k \partial Q''_l} \right)_0 \langle\, \psi'_v \mid Q''_k Q''_l \mid \psi''_v \,\rangle \end{aligned} \tag{1.30}$$

The overlap integral $\langle\, \psi'_v \mid \psi''_v \,\rangle$ is also referred as the Franck-Condon integral. Those in the second and third terms of equation 1.29 depend on the normal coordinates and therefore cannot be, in this form, generalized to any problem. To simplify the calculation of the integral given in equation 1.30, it is necessary to transform them in more versatile and usable Franck-Condon integrals.

For simplification purpose, the treatment of the Franck-Condon and overlap integrals will be done in the framework of the harmonic approximation. For the study of semi-rigid systems, where the Franck-Condon approximation is the most efficient, this approximation is in general sufficiently reliable and provides good results. In this framework, the total vibrational Hamiltonian, expressed in terms of normal coordinates and the associated momenta, can be written as a sum of one-dimensional vibrational Hamiltonians:

$$\mathcal{H}_v = \sum_{i=1}^{N} \mathcal{H}_{v_i} = \sum_{i=1}^{N} \mathcal{T}_i + \mathcal{V}_i = \sum_{i=1}^{N} -\frac{\hbar^2}{2} \frac{d^2}{dQ_i^2} + \frac{1}{2} \omega_i^2 Q_i^2 \tag{1.31}$$

where ω_i is the angular frequency of motion of the oscillator which is directly related to its frequency ν_i through the equation $\omega_i = 2\pi\nu_i$.

As noted in section 1.1, it is possible to write the vibrational wave function as a product of monodimensional wave functions:

$$\psi_v = \prod_{i=1}^{N} \psi_{v_i}(Q_i) \tag{1.32}$$

Each monodimensional wave function depends on one normal coordinate Q_i and one quantum number v_i. For clarity, the harmonic oscillators will be represented by their quantum numbers v_i. These vibrational quantum numbers are gathered in a vector $\boldsymbol{v}$ which defines unequivocally the vibrational state in the same way as the vibrational wave function.

Back to equation 1.29, the last two terms in its right-hand side can be developed as combinations of Franck-Condon integrals using the annihilation and creation operators, respectively a_i and $a_i^\dagger$. These operators are prime elements of the second quantization, a formalism commonly used to analyze many-body problems by shifting it to a matter involving one or two particles. To introduce them, we will consider the Schrödinger equation of the monodimensional harmonic oscillator. It is simple to connect our study to this particular case by recalling that in the harmonic approximation, the total energy of the isolated system can be written as the sum of N energies corresponding to each monodimensional oscillator, E_{v_i} :

$$E = \sum_{i=1}^{N} E_{v_i} \tag{1.33}$$

Applying the Hamiltonian operator, expressed as the sum of monodimensional Hamiltonian operators, to the wavefunction given in equation 1.32, we obtain N independent Schrödinger equations of one-dimensional harmonic oscillators:

$$\mathcal{H}_{v_i}\psi_{v_i} = E_{v_i}\psi_{v_i} \tag{1.34}$$

Using equation 1.31, the Schrödinger equation becomes:

$$-\frac{\hbar^2}{2}\frac{d^2\psi_{v_i}}{dQ_i^2} + \frac{1}{2}\omega_i^2 Q_i^2 \psi_{v_i} = E_{v_i}\psi_{v_i} \tag{1.35}$$

$$\frac{\hbar}{\omega_i}\frac{d^2\psi_{v_i}}{dQ_i^2} - \frac{\omega_i}{\hbar}Q_i^2\psi_{v_i} = -\frac{2E_{v_i}}{\omega\hbar}\psi_{v_i} \tag{1.36}$$

This equation can be simplified by replacing the normal coordinates $\mathbf{Q}$ by a new set of variables $\mathbf{q}$, using the following relation:

$$q_i = \sqrt{\frac{\omega_i}{\hbar}}\, Q_i \tag{1.37}$$

These variables are dimensionless and so are called *dimensionless normal coordinates*. By introducing them in equation 1.36, we obtain:

$$\frac{d^2\psi_{v_i}}{dq_i^2} + \left(\frac{2E_{v_i}}{\omega\hbar} - q_i^2\right)\psi_{v_i} = 0 \tag{1.38}$$

This equation can be solved exactly and its eigenvalues are the energies satisfying the relation

$$E_{v_i} = \hbar\omega_i\left(v_i + \frac{1}{2}\right) \tag{1.39}$$

where v_i is a quantum number belonging to the set of the nonnegative integers ($v_i = 0, 1, \ldots$). The eigenfunctions of equation 1.38 are given by the formula:

$$\psi_{v_i} = \mathcal{N}_{v_i} H_{v_i}(q_i)\, e^{-q_i^2/2} \tag{1.40}$$

where $\mathcal{N}_{v_i}$ is a normalization factor:

$$\mathcal{N}_{v_i} = \left(\frac{\omega_i}{\pi\hbar}\right)^{1/4} \frac{1}{\left[2^{v_i}(v_i!)\right]^{1/2}} \tag{1.41}$$

and $H_{v_i}(q_i)$ is a Hermite polynomial of degree v_i defined by:

$$H_{v_i}(q_i) = (-1)^{v_i} e^{q_i^2} \frac{d^{v_i}}{dq_i^{v_i}} e^{-q_i^2} \tag{1.42}$$

For any quantum number ($v_i + 1$ greater or equal to 2, it is possible to obtain the Hermite polynomial using the following recursion formula:

$$H_{v_i+1}(q_i) = 2q_i H_{v_i}(q_i) - 2v_i H_{v_i-1}(q_i) \tag{1.43}$$

We will now define the creation and annihilation operators mentioned previously as:

$$a_i^\dagger = \frac{1}{\sqrt{2}}\left(q_i - \frac{d}{dq_i}\right) \tag{1.44}$$

$$a_i = \frac{1}{\sqrt{2}}\left(q_i + \frac{d}{dq_i}\right) \tag{1.45}$$

If we apply the creation operator on equation 1.40, we obtain:

$$\begin{aligned} a_i^\dagger \psi_{v_i} &= \frac{1}{\sqrt{2}}\left(q_i - \frac{d}{dq_i}\right)\mathcal{N}_{v_i} H_{v_i}(q_i) e^{-q_i^2/2} \\ &= \frac{1}{\sqrt{2}} q_i \mathcal{N}_{v_i} H_{v_i}(q_i) e^{-q_i^2/2} - \frac{1}{\sqrt{2}}\mathcal{N}_{v_i}\left(\frac{dH_{v_i}(q_i)}{dq_i} e^{-q_i^2/2} + H_{v_i}(q_i)\frac{d\, e^{-q_i^2/2}}{dq_i}\right) \end{aligned} \tag{1.46}$$

Based on the definition of the Hermite polynomial given in equation 1.42 and using the relation given in equation 1.43 for the substitution, the derivative of the Hermite polynomial can be expressed as:

$$\frac{dH_{v_i}(q_i)}{dq_i} = 2q_i H_{v_i} - H_{v_i+1}(q_i) = 2v_i H_{v_i-1}(q_i) \tag{1.47}$$

Hence, equation 1.46 can be written

$$\begin{aligned} a_i^\dagger \psi_{v_i} &= \sqrt{2}\mathcal{N}_{v_i}\left(q_i H_{v_i}(q_i) - v_i H_{v_i-1}(q_i)\right) e^{-q_i^2/2} \\ &= \frac{1}{\sqrt{2}}\mathcal{N}_{v_i} H_{v_i+1}(q_i) e^{-q_i^2/2} \end{aligned} \tag{1.48}$$

From the equation of the nomalization factor 1.41, we can simply define that

$$\frac{1}{\sqrt{2}}\mathcal{N}_{v_i} = \left(\frac{\omega}{\pi\hbar}\right)^{1/4} \frac{(v_i+1)^{1/2}}{\left[2^{v_i+1}(v_i+1!)\right]^{1/2}} = (v_i+1)^{1/2}\mathcal{N}_{v_i+1} \tag{1.49}$$

As a consequence, equation 1.48 can be reduced to the following relation:

$$a_i^\dagger \psi_{v_i} = (v_i+1)^{1/2}\psi_{v_i+1} \tag{1.50}$$

Applied to the multidimensional vibrational wave function ψ_v, it is straightforward to see that the creation operator $a_i^\dagger$ will only apply to the one-dimensional wave function ψ_{v_i}. Using the notation of the vibrational quantum numbers $\boldsymbol{v}$, the effect of the creation operator on the vibrational wave function can be summed up as:

$$a_i^\dagger |\, \boldsymbol{v}'' \,\rangle = \sqrt{v_i''+1}\; |\, \boldsymbol{v}'' + 1_i'' \,\rangle \tag{1.51}$$

where $|\, \boldsymbol{v}''+1_i'' \rangle$ represents the vibrational state described by the multidimensional wave function $\psi''_{v+1_i} = \psi''_{v_1} \ldots \psi''_{v_{i-1}} \psi''_{(v_i+1)} \psi''_{v_{i+1}} \ldots \psi''_{v_N}$. From the energy perspective, this vibrational state differs from the state represented by $|\, \boldsymbol{v}'' \rangle$ by the eigenvalue of the monodimensional Schrödinger equation $E_{v_i''+1} = \hbar\omega_i''(v_i''+1/2) = E_{v_i''} + \hbar\omega_i''$.

Given this relation, we will represent the vibrational state whose wave function is ψ''_{v_i+1} in the compact form $|\, v_i'' + 1_i \,\rangle$ in our matrix notation.

In the same way, it is possible to show that the annihilation operator involves a "lowering" of the quantum numbers:

$$a_i |\, \boldsymbol{v}'' \,\rangle = \sqrt{v_i''}\; |\, \boldsymbol{v}'' - 1_i'' \,\rangle \tag{1.52}$$

From equations 1.44 and 1.45, the normal coordinate Q_i can be written

$$Q_i = \left(\frac{\hbar}{\omega_i}\right)^{1/2} q_i = \left(\frac{\hbar}{2\omega_i}\right)^{1/2} \left[a_i + a_i^\dagger\right] \tag{1.53}$$

Let us firstly consider the second term in the right hand side of equation 1.30, and more precisely the integral $\langle\, \boldsymbol{v}' \mid Q_k'' \mid \boldsymbol{v}'' \,\rangle$. We replace first the normal coordinate Q_k by the creation and annihilation operators using the relation given in equation 1.53. Applying these operators on the corresponding vibrational state as shown in equations 1.51 and 1.52, it is possible to develop this term in a combinations of Franck-Condon integrals.

$$\langle\, \boldsymbol{v}' \mid Q_k'' \mid \boldsymbol{v}'' \,\rangle = \langle\, \boldsymbol{v}' \mid \sqrt{\frac{\hbar}{2\omega_k''}}\; \left[a_i + a_i^\dagger\right] \mid \boldsymbol{v}'' \,\rangle \tag{1.54}$$

$$= \sqrt{\frac{\hbar}{2\omega_k''}}\; \left[\langle\, \boldsymbol{v}' \mid a_i \mid \boldsymbol{v}'' \,\rangle + \langle\, \boldsymbol{v}' \mid a_i^\dagger \mid \boldsymbol{v}'' \,\rangle\right] \tag{1.55}$$

$$= \sqrt{\frac{\hbar}{2\omega_k''}}\; \left[\sqrt{v_k''}\, \langle\, \boldsymbol{v} \mid \boldsymbol{v}'' - 1_k'' \,\rangle + \sqrt{v_k''+1}\, \langle\, \boldsymbol{v} \mid \boldsymbol{v}'' + 1_k'' \,\rangle\right] \tag{1.56}$$

Following the same approach, it is possible to develop $\langle\, \boldsymbol{v}' \mid Q_k'' Q_l'' \mid \boldsymbol{v}'' \,\rangle$ in a combination of

Franck-Condon integrals. We present firstly the case $l \neq k$, more straightforward to formulate.

$$\langle \boldsymbol{v}' \mid Q''_k Q''_l \mid \boldsymbol{v}'' \rangle = \langle \boldsymbol{v}' \mid \sqrt{\frac{\hbar}{2\omega''_k}} \left[a_i + a_i^\dagger\right] \sqrt{\frac{\hbar}{2\omega''_l}} \left[a_j + a_j^\dagger\right] \mid \boldsymbol{v}'' \rangle \tag{1.57}$$

$$\begin{aligned} &= \frac{\hbar}{2\left(\sqrt{\omega}''_k \sqrt{\omega}''_l\right)} \Big[\langle \boldsymbol{v}' \mid a_i a_j \mid \boldsymbol{v}'' \rangle + \langle \boldsymbol{v}' \mid a_i a_j^\dagger \mid \boldsymbol{v}'' \rangle \\ &\quad + \langle \boldsymbol{v}' \mid a_i^\dagger a_j \mid \boldsymbol{v}'' \rangle + \langle \boldsymbol{v}' \mid a_i^\dagger a_j^\dagger \mid \boldsymbol{v}'' \rangle\Big] \end{aligned} \tag{1.58}$$

$$\begin{aligned} &= \frac{\hbar}{2\left(\sqrt{\omega}''_k \sqrt{\omega}''_l\right)} \Big[\sqrt{v''_k v''_l}\, \langle \boldsymbol{v}' \mid \boldsymbol{v}'' - 1''_k - 1''_l \rangle \\ &\quad + \sqrt{v''_k (v''_l + 1)}\, \langle \boldsymbol{v}' \mid \boldsymbol{v}'' - 1''_k + 1''_l \rangle \\ &\quad + \sqrt{(v''_k + 1) v''_l}\, \langle \boldsymbol{v}' \mid \boldsymbol{v}'' + 1''_k - 1''_l \rangle \\ &\quad + \sqrt{(v''_k + 1)(v''_l + 1)}\, \langle \boldsymbol{v}' \mid \boldsymbol{v}'' + 1''_k + 1''_l \rangle\Big] \end{aligned} \tag{1.59}$$

In the particular case $l = k$, the annihilation and creation operators are applied on the same mode and the resulting expression of $\langle \boldsymbol{v}' \mid Q''^2_k \mid \boldsymbol{v}'' \rangle$ can be written:

$$\langle \boldsymbol{v}' \mid Q''^2_k \mid \boldsymbol{v}'' \rangle = \frac{\hbar}{2\omega''_k} \langle \boldsymbol{v}' \mid a_i a_i + a_i a_i^\dagger + a_i^\dagger a_i + a_i^\dagger a_i^\dagger \mid \boldsymbol{v}'' \rangle \tag{1.60}$$

$$\begin{aligned} &= \frac{\hbar}{2\omega''_k} \Big[\sqrt{v''_k (v''_k - 1)}\, \langle \boldsymbol{v}' \mid \boldsymbol{v}'' - 2''_k \rangle \\ &\quad + (2v''_k + 1)\, \langle \boldsymbol{v}' \mid \boldsymbol{v}'' \rangle \\ &\quad + \sqrt{(v''_k + 1)(v''_k + 2)}\, \langle \boldsymbol{v}' \mid \boldsymbol{v}'' + 2''_k \rangle\Big] \end{aligned} \tag{1.61}$$

The complete formula of the transition dipole moment integral including equations 1.56, 1.59 and 1.61 can be found in appendix A.

1.4 A particular case of the Franck-Condon integrals : $\langle \mathbf{0}' \mid \mathbf{0}'' \rangle$

In the previous section, the transition dipole moment integral has been expressed with respect to the electronic transition moment and its derivatives as well as the Franck-Condon integrals. The electronic transition moment and its derivatives can be expressed with finite differences and their calculation can be performed by some quantum mechanical methods such as *Time Dependent Density Functional Theory*. We will consider for now that they are already known and will focus in this section on the calculation of the Franck-Condon integrals.

As a starting point, we will consider in this section the overlap integral between the vibrational ground states of both electronic states involved in the transition, $\langle \mathbf{0}' \mid \mathbf{0}'' \rangle$. Each vibrational wave function is the product of one-dimensional wave functions, solutions of the Schrödinger equation of the harmonic oscillator given in equation 1.38, so that we can write, in normal coordinates:

$$\mid \mathbf{0} \rangle = \prod_{i=1}^{N} \mathcal{N}_{(v_i=0)} e^{-\omega_i Q_i^2/2\hbar} = \prod_{i=1}^{N} \left(\frac{\omega_i}{\pi\hbar}\right)^{1/4} e^{-\omega_i Q_i^2/2\hbar} \tag{1.62}$$

To simplify the previous equation, we will introduce a new variable $\gamma_i = \omega_i/\hbar$, that will be called the reduced frequency. Consequently, the vibrational wave function has the following expression

$$| \mathbf{0} \rangle = \prod_{i=1}^{N} \left(\frac{\gamma_i}{\pi} \right)^{1/4} e^{-\gamma_i Q_i^2/2} \tag{1.63}$$

From equation 1.63, it can be seen that a problem will arise when trying to calculate the overlap integral, as each vibrational wave function is expressed in a different set of normal coordinates. Hence, a way to express a set of normal coordinates with respect to another one will be convenient. However, in most cases, the transformation is not necessarily obvious and the mixing of the normal modes during the electronic transition needs to be taken into account. Duschinsky [31] took an early interest in this problem when studying polyatomic molecules. He proposed a linear transformation between the normal coordinates of both states to deal with the mixing of the normal modes in the transition, often called Duschinsky effect in the literature. In our case, we will express the normal coordinates of the initial state as a linear combination of the normal coordinates of the final state:

$$\mathbf{Q}' = \mathbf{J}\mathbf{Q}'' + \mathbf{K} \tag{1.64}$$

This transformation is a good approximation when the molecule does not undergo a noticeable distortion during the transition. An extensive discussion of the transformation between the normal coordinates including a possible distortion of the molecule and the limitations of the Duschinsky transformation has been done by Lucas [32]. This case is rather rare for medium and large molecules where the conditions described in section 1.3 to apply the Franck-Condon principle are met. As a consequence, we will limit ourselves to the Duschinsky transformation to express one set of normal coordinates with respect to the other one.

This linear transformation can be straightforwardly obtained. Let us consider a shift vector $\mathbf{S}$ of the nuclear Cartesian coordinates between the initial and final states. Using this vector, we can express the nuclear coordinates of the initial state with respect to those of the final state:

$$\mathbf{X}'_{\text{eq}} = \mathbf{X}''_{\text{eq}} + \mathbf{S} \tag{1.65}$$

For simplification, we will transpose this relation to the mass-weighted Cartesian coordinates and we will use the more common notation $\mathbf{L}^{-1}$ to represent the pseudoinverse matrix of $\mathbf{L}$. We can now switch to normal coordinates thanks to the relation given in equation 1.28 .

$$\mathbf{L}'\mathbf{Q}' = \mathbf{L}''\mathbf{Q}'' + \mathbf{M}^{1/2}\mathbf{S} \tag{1.66}$$

The relation between normal coordinates of the initial and final states can be written:

$$\mathbf{Q}' = \left(\mathbf{L}'\right)^{-1}\mathbf{L}''\mathbf{Q}'' + \left(\mathbf{L}'\right)^{-1}\mathbf{M}^{1/2}\mathbf{S} \tag{1.67}$$

Comparing equations 1.64 and 1.67, the rotation matrix and shift vectors can be defined with the following relations:

$$\mathbf{J} = \left(\mathbf{L}'\right)^{-1}\mathbf{L}'' \quad \text{and} \quad \mathbf{K} = \left(\mathbf{L}'\right)^{-1}\mathbf{M}^{1/2}\mathbf{S}$$

Because $\mathbf{Q}'$ and $\mathbf{Q}''$ form each one an orthonormalized basis set, the columns of $\mathbf{L}'$ and $\mathbf{L}''$ are orthonormalized. This makes easier the computation of $\mathbf{J}$ and $\mathbf{K}$ as the left inverse $\left(\mathbf{L}'\right)^{-1}$ can be replaced by the transpose of $\mathbf{L}'$.

$$\mathbf{J} = \left(\mathbf{L}'\right)^{\mathrm{T}}\mathbf{L}'' \quad \text{and} \quad \mathbf{K} = \left(\mathbf{L}'\right)^{\mathrm{T}}\mathbf{M}^{1/2}\mathbf{S}$$

A consequence of this expression and the orthonormalization of the columns of $\mathbf{L}'$ and $\mathbf{L}''$ is that $\mathbf{J}$ is an orthogonal matrix. We can directly deduce two interesting properties for $\mathbf{J}$, the inverse of $\mathbf{J}$ can be replaced by its transpose, which does not need calculation, and the sum of the square of each element in a row or a column of $\mathbf{J}$ is equal to unity.

Based on this transformation, it is now possible to express both vibrational wave function in the same set of normal coordinates $\mathbf{Q}''$. Since $\psi'_{v=0}$ is expanded in terms of an orthonormalized set $\mathbf{Q}''$, we need to control that the overlap integrals are still normalized with this change. This can be easily done by checking that $\psi'_{v=0}$ is appropriately normalized :

$$\mathcal{L}^2 \int |\psi_v|^2 d\mathbf{Q}'' = 1 \tag{1.68}$$

where $\mathcal{L}$ is the normalization constant. Applying $\mathbf{J}^{-1}$ in equation 1.64, we obtain the "inverse" Duschinsky transformation to express $\mathbf{Q}''$ with respect to $\mathbf{Q}'$

$$\mathcal{L}^2 \int |\psi_v|^2 d(\mathbf{J}^{-1}\mathbf{Q}' + \mathbf{J}^{-1}\mathbf{K}) = \mathcal{L}^2 \int |\psi_v|^2 \mathbf{J}^{-1} d\mathbf{Q}' = 1 \tag{1.69}$$

Therefore, the normalization constant $\mathcal{L}$ is equal to the square root of the Jacobian determinant, $\det(\mathbf{J})^{1/2}$. The normalized overlap integral between the fundamental vibrational states can be computed with the inclusion of this constant.

$$\begin{aligned}\langle\, \mathbf{0}' \mid \mathbf{0}'' \,\rangle = \det(\mathbf{J})^{1/2} &\left(\frac{\gamma'_1}{\pi}\right)^{1/4} \dots \left(\frac{\gamma'_N}{\pi}\right)^{1/4} \left(\frac{\gamma''_1}{\pi}\right)^{1/4} \dots \left(\frac{\gamma''_N}{\pi}\right)^{1/4} \\ &\int_{-\infty}^{+\infty} dQ''_1 \dots dQ''_N e^{-\gamma'_1 {Q'_1}^2/2} \dots e^{-\gamma'_N {Q'_N}^2/2} e^{-\gamma''_1 {Q''_1}^2/2} \dots e^{-\gamma''_N {Q''_N}^2/2}\end{aligned} \tag{1.70}$$

To simplify the above equation, we will switch to a more compact matrix notation.

$$\langle\, \mathbf{0}' \mid \mathbf{0}'' \,\rangle = \pi^{-N/2} \det\left[\mathbf{\Gamma}' \mathbf{\Gamma}''\right]^{1/4} \det(\mathbf{J})^{1/2} \int_{-\infty}^{+\infty} dQ''_1 \dots dQ''_N e^{-\frac{1}{2}(\mathbf{Q}'^{\mathrm{T}}\mathbf{\Gamma}'\mathbf{Q}' + \mathbf{Q}''^{\mathrm{T}}\mathbf{\Gamma}''\mathbf{Q}'')} \tag{1.71}$$

where $\mathbf{\Gamma}'$ and $\mathbf{\Gamma}''$ are diagonal matrices of the reduced frequencies γ'_i and γ''_i, respectively. To calculate the integral in equation 1.71, it is preferable to proceed to a change of variables. The details of this change are given in appendix B. We will only give its main features to obtain the overlap integral.

We introduce a new vector $\mathbf{a}$ defined as a linear transformation of the normal coordinates $\mathbf{Q}''$:

$$\mathbf{a} = \mathbf{b}\mathbf{Q}'' + \mathbf{c} \tag{1.72}$$

This vector is chosen so that the following identity is satisfied:

$$\mathbf{Q'}^{\mathrm{T}}\mathbf{\Gamma'}\mathbf{Q'} + \mathbf{Q''}^{\mathrm{T}}\mathbf{\Gamma''}\mathbf{Q''} = \mathbf{a}^{\mathrm{T}}\mathbf{a} + \mathbf{d} \tag{1.73}$$

The interest of this procedure is to easily calculate the integration over the whole configuration space by using the mathematical result

$$\int_{-\infty}^{+\infty} e^{-\lambda x^2} dx = \sqrt{\frac{\pi}{\lambda}} \tag{1.74}$$

By equating the coefficients of $\mathbf{Q''}$ in equation 1.73, we obtain the following results for $\mathbf{b}$, $\mathbf{c}$ and $\mathbf{d}$:

$$\begin{aligned}
\mathbf{b} &= (\mathbf{J}^{\mathrm{T}}\mathbf{\Gamma'}\mathbf{J} + \mathbf{\Gamma''})^{1/2} \\
\mathbf{c} &= (\mathbf{J}^{\mathrm{T}}\mathbf{\Gamma'}\mathbf{J} + \mathbf{\Gamma''})^{-1/2}\mathbf{J}^{\mathrm{T}}\mathbf{\Gamma'}\mathbf{K} \\
\mathbf{d} &= \mathbf{K}^{\mathrm{T}}\mathbf{\Gamma'}\mathbf{K} - \mathbf{K}^{\mathrm{T}}\mathbf{\Gamma'}\mathbf{J}(\mathbf{J}^{\mathrm{T}}\mathbf{\Gamma'}\mathbf{J} + \mathbf{\Gamma''})^{-1}\mathbf{J}^{\mathrm{T}}\mathbf{\Gamma'}\mathbf{K}
\end{aligned}$$

Changing the variables from $\mathbf{Q''}$ to $\mathbf{a}$, and replacing $\mathbf{b}$ and $\mathbf{d}$ by their value, the overlap integral $\langle\, \mathbf{0'} \mid \mathbf{0''} \,\rangle$ can be written:

$$\begin{aligned}
\langle\, \mathbf{0'} \mid \mathbf{0''} \,\rangle = \pi^{-N/2} \det\left[\mathbf{\Gamma'}\,\mathbf{\Gamma''}\right]^{1/4} \det(\mathbf{J})^{1/2} e^{\left[-\frac{1}{2}\mathbf{K}^{\mathrm{T}}\mathbf{\Gamma'}\mathbf{K} + \frac{1}{2}\mathbf{K}^{\mathrm{T}}\mathbf{\Gamma'}\mathbf{J}(\mathbf{J}^{\mathrm{T}}\mathbf{\Gamma'}\mathbf{J}+\mathbf{\Gamma''})^{-1}\mathbf{J}^{\mathrm{T}}\mathbf{\Gamma'}\mathbf{K}\right]} \\
\times \frac{1}{\det\left[\mathbf{J}^{\mathrm{T}}\mathbf{\Gamma'}\mathbf{J} + \mathbf{\Gamma''}\right]^{1/2}} \int_{-\infty}^{+\infty} d\mathrm{a}_1 \ldots d\mathrm{a}_N e^{-\frac{1}{2}\mathbf{a}^{\mathrm{T}}\mathbf{a}}
\end{aligned} \tag{1.75}$$

As a result, the normalized overlap integral between each vibrational ground state of the electronic states is given by the following equation:

$$\begin{aligned}
\langle\, \mathbf{0'} \mid \mathbf{0''} \,\rangle = 2^{N/2} \det\left[\mathbf{\Gamma'}\,\mathbf{\Gamma''}\right]^{1/4} \left[\frac{\det(\mathbf{J})}{\det\left(\mathbf{J}^{\mathrm{T}}\mathbf{\Gamma'}\mathbf{J} + \mathbf{\Gamma''}\right)}\right]^{1/2} \\
\times \exp\left[-\frac{1}{2}\mathbf{K}^{\mathrm{T}}\mathbf{\Gamma'}\mathbf{K} + \frac{1}{2}\mathbf{K}^{\mathrm{T}}\mathbf{\Gamma'}\mathbf{J}(\mathbf{J}^{\mathrm{T}}\mathbf{\Gamma'}\mathbf{J} + \mathbf{\Gamma''})^{-1}\mathbf{J}^{\mathrm{T}}\mathbf{\Gamma'}\mathbf{K}\right]
\end{aligned} \tag{1.76}$$

Chapter 2

Strategies to compute the Franck-Condon integral $\langle\, \boldsymbol{v}' \mid \boldsymbol{v}'' \,\rangle$

2.1 Introduction

In this chapter, we will consider the general case $\langle\, \boldsymbol{v}' \mid \boldsymbol{v}'' \,\rangle$ where $\boldsymbol{v}'$ and $\boldsymbol{v}''$ are non-fundamental vibrational states. Mathematically, we will suppose $\mid \boldsymbol{v}' \,\rangle \neq \mid \mathbf{0} \,\rangle$ and $\mid \boldsymbol{v}'' \,\rangle \neq \mid \mathbf{0} \,\rangle$[a]. Using equation 1.40, the overlap integral is given, in matrix notation, by the relation

$$\begin{aligned}\langle\, \boldsymbol{v}' \mid \boldsymbol{v}'' \,\rangle = \pi^{-N/2} \det\left[\boldsymbol{\Gamma}' \boldsymbol{\Gamma}''\right]^{1/4} \det(\mathbf{J})^{1/2} \left(\prod_{i=1}^{N} 2^{v_i'} 2^{v_i''} (v_i'!)(v_i''!)\right)^{-1/2} \\ \times \int_{-\infty}^{+\infty} dQ_1'' \dots dQ_N'' \left(\prod_{i=1}^{N} H_{v_i'}(\Gamma_i'^{1/2} Q_i') H_{v_i''}(\Gamma_i''^{1/2} Q_i'')\right) \\ \times \exp\Big[-\frac{1}{2}(\mathbf{Q}'^{\mathrm{T}} \boldsymbol{\Gamma}' \mathbf{Q}' + \mathbf{Q}''^{\mathrm{T}} \boldsymbol{\Gamma}'' \mathbf{Q}'')\Big]\end{aligned} \tag{2.1}$$

From the definition of the Hermite polynomial given in equation 1.42, it is plain to see that the overlap integral defined above cannot be easily evaluated. Hence, a direct calculation is generally out of question and another, more accessible procedure needs to be devised.

Without pretending to exhaustivity, we will present several approaches to the calculation of the overlap integrals. These methods are sometimes intended for particular cases or tackle the problem from different perspectives, and it may be sometimes difficult to find out which one is the best, as we will be show below. Hence, we will favor a chronological presentation of these methods and focus afterward on those used in this work.

Earlier works, mainly centered on diatomic molecules, were presented soon after Franck stated the principle of a slow nuclear motion with respect to the electronic transition [16] and Condon discussed the mathematical formalism of this theory [17,33] to apply it to bands study. Hutchisson, in 1930, proposed an analytic method to calculate the overlap integrals of diatomic molecules using the harmonic approximation [34]. In this case, restricted to linear oscillations, only monodimensional oscillators are involved in the electronic transition. Hutchisson developed the Hermitian polynomial using a generating function and was able to formulate an analytic expression of the overlap integrals

[a] The formalism presented here is also applicable to the case where one of the vibrational state is fundamental

to study the main bands of Na_2, K_2, I_2 and H_2. Later, he removed the restriction on the linear oscillation, taking into account the anharmonicity of the oscillator with the Schrödinger theory of perturbation [35]. A few decades later, Manneback [36] proposed a recursive approach to solve the overlap integrals of diatomic molecules. This method was seemingly less general than Hutchisson's as it was devised for the case of harmonic one-dimensional oscillators. Smith [37] showed that recursion formulae could be directly derived from Hutchisson's analytic expressions of the overlap integrals. At the same time, new methods started to favor the less cumbersome second quantization to express the transition dipole moment integral instead of the analytic formulae. Although it had been firstly used to analyze the one-dimensional harmonic oscillator [38, 39], the formalism of the second quantization is of even greater importance when dealing with multidimensional harmonic oscillators where the complete formula of the overlap integral becomes really complex and difficult to read.

At this same period also a thorough development of theoretical methods to formulate the Franck-Condon integrals in the case of polyatomic molecules really started, almost three decades after Herzberg and Teller had first discussed the application of the Franck principle to polyatomic molecules and the resulting selection rules [29]. A first example of polyatomic molecule was studied in 1950 by Craig [40] but the author restricted its analysis on the totally symmetrical vibration of benzene. Coon et al. [41] used the Franck-Condon principle to study the structure of polyatomic molecules in excited state from the vibronic bands of ultraviolet absorption spectra. Having extended the application of analytic formulae to more generic systems, they still restricted them to symmetrical vibrations. In their original work, they focused on triatomic systems with a C_{2v} symmetry where few possible overlap integrals need to be taken into account, but similar approaches were also applied to larger systems [42].

Sharp and Rosenstock [43] used generating functions to bypass the problematic calculation of the Hermite polynomials, equating the coefficients of dummy variables, and were able to obtain general formulae for the Franck-Condon integrals of polyatomic molecules. The calculations were set in the framework of the harmonic approximation but did not impose any restriction on the symmetry of the vibration, devising a general-purpose method. This method can be seen as a generalization of the one presented by Hutchisson and we will present it in greater details further, as the formalism they introduced underlie our calculations of the transition dipole moment integrals. The original published work [43] contained a listing of overlap integrals for transitions to overtons and combination levels up to four simultaneously excited modes and the equations took into account excited initial states for the transitions (hotbands). However, the quickly-raising complexity of the analytic formulae were error-prone and many works afterwards preferred, partly for this reason, to use recursive formulae. Correction of these original formulae have been proposed later [44], sometimes accompanied by extensions of the set of analytic formulae [45].

Overcoming the previous restrictions on the applicable systems to perform the Franck-Condon calculations, the original work of Sharp and Rosenstock found a large audience and was used as a direct application [46] or as a basis for new developments in analytic methods. Baranov *et al.* [47,48] used the generating functions, among other models, to obtain general expressions to evaluate any overlap integrals with up to two excited normal vibrations in each vibronic state. Their approach breaks down two remaining limitations in the original analytic formulae. The first one was purely conceptual and concerned the approximation of the electronic transition dipole moment. While

Sharp and Rosenstock restricted their study to the Franck-Condon approximation, Baranov *et al.* extended it to the Herzberg-Teller approximation. Weber and Hohlneicher [45], among others, showed that the original analytic formulae were compatible with higher order of the Taylor series of the electronic transition dipole moment. The other limitation concerned the quantum numbers of each excited mode. The formulae obtained by Baranov *et al.* [47] could be used directly for any combination while restricted to two excited modes. To simulate the photoelectron spectrum of dichlorocarbene, Kohn *et al.* [49] used the formulae of the overlap integrals obtained with the Sharp and Rosenstock method to devise a general expression for the cases with the final vibrational state corresponding to combination bands of three simultaneously excited modes. Contrary to Baranov *et al.*, they only considered transitions from the fundamental vibrational state. While not using the formalism of Sharp and Rosenstock, Mebel *et al.* [50, 51] contributed to the generalization of the analytic formula treating the case of four simultaneously excited modes in the final state. Finally, Islampour *et al.* [52] proposed a general closed-form expression with no restriction on the number of simultaneously excited modes in the initial and final states. Kikuchi *et al.* [53] proposed an alternative expression stemming directly from the generating function given by Sharp and Rosenstock. On the whole, the complexity of these more general formulae, caused by the piling of sums and products, renders difficult their implementation in a general-purpose optimized procedure. Recently, Chang [54, 55] proposed new and simpler formulae for overlap integrals with up to four excited modes in the final state. However, they are limited to transitions from the fundamental vibrational state. Other approaches [56, 57] were also employed to obtain analytic formulae. Faulkner and Richardson [56] proposed two methods for calculating the transition dipole moment integrals, at the level of approximation of Franck-Condon, as well as Herzberg-Teller. Their first method consisted of a linear transformation of the normal coordinates in the initial and final state to remove the Duschinsky rotation and so be able to express the general overlap integrals as a product of monodimensional integrals. While their original work assumed that at least one of the vibrational state involved in the transitions was the ground state, Kulander [58] removed this restriction. Their second method is based on a perturbative expansion of the vibrational wave functions of the excited electronic state in terms of the ground electronic state vibrational wave functions. However, as acknowledged by the authors, the slow convergence of the perturbative theory expressions mostly ruled it out as an efficient general-purpose procedure in actual computation. Recently, Borrelli and Peluso [59] proposed a perturbative method to handle the normal mode mixing and calculate the Franck-Condon integrals. Their method offers to reduce the storage requirement for the calculations but at the expense of increasing computational costs when dealing with large systems.

The difficulty to generate straightforward analytic expressions and encode them into a automatic procedure to compute the transition dipole moment integrals make often more interesting the use of recursion instead. Ruhoff [60,61] proposed a complete recursive approach based on the generating formula of Sharp and Rosenstock by generalizing Lermé's [62] procedure for two-dimensional Franck-Condon factors. Since it is using exactly the same formalism, this method can be implemented in a program as a unique way to calculate the Franck-Condon integrals or together with analytic formulae as a "back-up" when the latter cannot cover all necessary transitions. Previously, Kupka and Cribb [63] used a different approach to obtain generating functions similar to Sharp and Rosenstock. From them, recursion expressions were then derived and could also be used in

addition to analytic ones.

Slightly later than Sharp and Rosenstock, Doktorov *et al.* [64, 65] used the coherent states described by Glauber [66] to derive recurrence relations for any overlap integrals without regard to the number of simultaneously excited modes in each state. Their work, providing an out-of-the-box method theoretically applicable to any molecular system, has been widely used in Franck-Condon calculations program [67–70]. The equivalence of the formulae obtained by Doktorov *et al.* and Sharp and Rosenstock to calculate the overlap integrals has been shown by Liang and Haiyang [71]. More recently, Malmqvist amd Forsberg [72] used LU decomposition procedures[b] to obtain triangular matrices and derive recursion formulae to evaluate the overlap integrals for multidimensional harmonic oscillators.

When dealing with polyatomic molecules, most works were done in the framework of the harmonic approximation. While most methods are general assumed to be valid in the case of anharmonicity, the adaptation is rarely discussed. There are several reasons to this neglect. A simple one lies on the complexity of the expressions with the "simple" case of the harmonic oscillator. Developing an efficient and versatile method to compute the transition intensities and generate the vibronically resolved spectrum is not straightforward. For this reason, it is generally advisable to focus on the harmonic case. Another reason concerns the speed of the calculations. Even simple schemes to introduce the anharmonicity cause an important increase in the number of terms in the equations of the Franck-Condon integrals and in the calculations of the electronic structures. Hazra *et al.* [73] proposed a simple scheme to account partially for the anharmonicity, aiming primarily at resolving the failures of the vertical Franck-Condon approach. They expanded the final state potential energy surface around the equilibrium geometry of the initial state assuming the separability of the normal modes. In this case, it might happen that for some modes, the inappropriate curvature of the potential energy surface leads to imaginary frequencies. If a frequency is real, then the potential is treated with the harmonic approximation. Otherwise, they calculate the corresponding complete one-dimensional potential at the anharmonic level to construct the excited-state vibrational Hamiltonian. Luis *et al.* [74, 75] proposed a method to simulate a vibronically resolved spectrum taking into account the anharmonicity. The latter is introduced as a second-order perturbation treatment of the Franck-Condon factors initially calculated at the harmonic level. Because of the computational costs, their study remained at the scale of small systems.

It is noteworthy that in our work, we remained in the framework of the harmonic approximation.

2.2 Analytic formulae and recursive methods

Based on the methods we have just presented, we can identify two general approaches of the problem of calculation of the Franck-Condon integrals. The recursive and analytic approaches are not restricted to Franck-Condon calculations but represent generic ways to deal with a matter where multiple calculations, often interlinked, are required. Because each one has its own advantages and drawbacks, a good knowledge of their applicability and their limits is necessary to take fully ad-

[b] A procedure for decomposing an $n \times n$ matrix $\mathbf{A}$ into a product of a lower triangular matrix $\mathbf{L}$ and an upper triangular matrix $\mathbf{U}$

vantage of each one. Also, because calculations are now mostly, if not all, performed by computers, we must also take into account the implication of the choice of one method with respect to the efficiency of the program and the resources it will need.

Analytic methods rely on a set of formulae to perform a calculation. They often are a good choice when the problem to handle is relatively small or can be sufficiently simplified. Their underlying principle is to have a set of specialized equations for each task. In our case, the calculation of overlap integrals, the generating functions used in the most widely adopted analytic approaches will give a formula for each kind of transition, with a precise set of quantum numbers representing each electronic state. Having a large number of equations, each one for a transition, enables a quick computation of the Franck-Condon integrals. However, because these equations are complex, they must be generated beforehand. As a direct consequence, a first, major restriction of these analytic approaches is the need to generate all the required overlap integrals before running the actual calculation and store them in the core of the program. Hence, this strategy induces a marked growth of the program size. Also, because all formulae need to be created beforehand, one needs to "predict" the possible transitions, so it is difficult to design a general-purpose program based solely on this approach. This limitation is not a problem when analyzing small systems and some particular, symmetrical, medium size systems. However, for large systems the number of possible combinations of excited modes with their respective quantum numbers increase steeply as the number of normal vibrations grow, so that it is unrealistic to consider the whole of the transitions. Even for other analytic methods, such as those based on closed-form expressions [47, 49, 50, 52, 53], the piling of loops means a drop in performances and their calculations lose rapidly in efficiency. Consequently, even these methods are difficult to consider for a really versatile use. While they can theoretically be used without regard to the value of the quantum numbers, the most advanced of these methods consider up to four modes, the formulae of the overlap integrals being already really complex.

Nevertheless, the greatest advantage of these methods is the easiness and efficiency to compute the overlap integrals once they have been formulated, along with their accuracy. The program will not need much of the memory of the computer to run but many accesses to different equations to compute a given spectrum. Another grievance made to the analytic approach is the risk of errors in the formulae, which might be difficult to spot *a posteriori*. However, this risk is now generally low because the calculations are treated by computational tools dedicated to this kind of tasks.

On the other side, the recursive approach only needs a few, often simple, formulae. The underlying concept, applied to Franck-Condon integrals, is that when considering some transition, it is likely that similar transitions have been calculated before, with a few differences in the vibrational states. Ideally, the previously treated transition differs from the current one by only one quantum number altogether. Based on this idea, each overlap integral could be directly computed from already calculated integrals with the same formula or restricted set of formulae. A direct consequence is that the code will not be weighed down by a large number of equations to store in the code. Likewise, because any new integral is connected to previously calculated ones, it means that only the overlap integrals between ground states need to be known exactly and any other transition can be evaluated from it. This is a notable advantage of the recursion approach with respect to the analytic one, which makes it perfectly adapted to large molecules. However, this interconnection

between overlap integrals can become really cumbersome when treating large systems with many possible transitions to deal with. Because any new overlap integrals need one-to-many previously calculated overlap integrals, the latter need to be stored. Without entering into the technical implications of the storage that will be discussed in section 2.6, the choice made at this point in the program can be really costly for the efficiency of the code. Hence, recursion approaches always need to find a good compromise between storage and re-calculation, some overlap integrals not being stored and rather recalculated on demand in the latter case. Finally, because the computations rely on previously calculated Franck-Condon integrals, the risk of a propagation of error exists. However, this risk is limited by the low number of formulae -and so the transcription errors- and the reachable precision of the computers nowadays which should be far higher than needed. However, a strict control might be required when dealing with large sequences of recursive calculations to ensure that the imprecisions of the results cannot pile up to a non-negligible error.

To summarize the main advantages of both methods, we can say that on a controlled number of combinations, analytic formulae give quick and precise results while a recursive approach is needed for a general-purpose method applicable to large molecules. We will now present in the following sections two approaches, the original Sharp-Rosenstock [43] analytic approach with the generating function and a recursive derivation from it done by Ruhoff [60].

2.3 The general analytic approach of Sharp and Rosenstock

As presented before, Sharp and Rosenstock proposed early a method to compute any kind of overlap integrals without restriction on the nature of the transition. This method has the advantage to be straightforward, sometimes at the cost of the compactness of the formulae. Because the method is general enough to consider any kind of transition, whatever the number of simulaneously excited modes in the initial and final states, it is still widely used and referred to. We will present here its principle.

The main idea of the method designed by Sharp and Rosenstock is to use the overlap integrals $\langle\, \boldsymbol{v}' \mid \boldsymbol{v}'' \,\rangle$ as coefficients in a power series of dummy vectors $\mathbf{T}$ and $\mathbf{U}$. Formally, this power series should be written:

$$f(\mathbf{T},\mathbf{U}) = \sum_{v'_1,\ldots,v'_N=0}^{\infty} \sum_{v''_1,\ldots,v''_N=0}^{\infty} \prod_{i=1}^{N} \mathrm{T}_i^{v'_i} \prod_{j=1}^{N} \mathrm{U}_j^{v''_j} \left(\prod_{i=1}^{N} \frac{2^{v'_i} 2^{v''_i}}{v'_i! v''_i!} \right)^{1/2} \langle\, \boldsymbol{v}' \mid \boldsymbol{v}'' \,\rangle \tag{2.2}$$

where $\sum_{v''_1,\ldots,v''_N=0}^{\infty}$ represents N sums, one for each quantum number v_i.

This formal mathematical notation is really cumbersome to deal with while developing the power series. To ease the reading of the following equations, we will adopt a more compact albeit less correct pure matrix notation. In this notation, the previous equation can be simply written:

$$f(\mathbf{T},\mathbf{U}) = \sum_{\boldsymbol{v}'} \sum_{\boldsymbol{v}''} \mathbf{T}^{\boldsymbol{v}'} \mathbf{U}^{\boldsymbol{v}''} \left(\frac{2^{\boldsymbol{v}'} 2^{\boldsymbol{v}''}}{\boldsymbol{v}'! \boldsymbol{v}''!} \right)^{1/2} \langle\, \boldsymbol{v}' \mid \boldsymbol{v}'' \,\rangle \tag{2.3}$$

Replacing the overlap integral by its value given in equation 2.1, the power series can be ex-

panded as the following sum:

$$
\begin{aligned}
f(\mathbf{T},\mathbf{U}) = \frac{\pi^{-N/2}}{\boldsymbol{v}'!\boldsymbol{v}''!}\det\left[\boldsymbol{\Gamma}'\,\boldsymbol{\Gamma}''\right]^{1/4}\det(\mathbf{J})^{1/2}\sum_{\boldsymbol{v}'}\sum_{\boldsymbol{v}''}\mathbf{T}^{\boldsymbol{v}'}\mathbf{U}^{\boldsymbol{v}''} \\
\times\int_{-\infty}^{+\infty} d\mathbf{Q}''\left(H_{v'}(\boldsymbol{\Gamma}'^{1/2}\mathbf{Q}')H_{v''}(\boldsymbol{\Gamma}''^{1/2}\mathbf{Q}'')\right)\exp\left[-\frac{1}{2}(\mathbf{Q}'^{\mathrm{T}}\boldsymbol{\Gamma}'\mathbf{Q}'+\mathbf{Q}''^{\mathrm{T}}\boldsymbol{\Gamma}''\mathbf{Q}'')\right]
\end{aligned}
\tag{2.4}
$$

We will define the generating function of the Hermite polynomial [76] $H_{v''}(\boldsymbol{\Gamma}''^{1/2}\mathbf{Q}'')$ in matrix notation with the following identity

$$
\sum_{\boldsymbol{v}''=0}^{\infty}\frac{\mathbf{U}^{\boldsymbol{v}''}}{\boldsymbol{v}''!}H_{v''}(\boldsymbol{\Gamma}''^{1/2}\mathbf{Q}'') = \exp\left[-\mathbf{U}^2+2\mathbf{U}^{\mathrm{T}}\boldsymbol{\Gamma}''^{1/2}\mathbf{Q}''\right] \tag{2.5}
$$

A similar equation can be obtained with the Hermite polynomial $H_{v'}(\boldsymbol{\Gamma}'^{1/2}\mathbf{Q}')$ and the dummy variable $\mathbf{T}$. Similarly to the problem described while studying the overlap integrals between vibrational ground states, we need to transform the normal coordinates of the initial states using the Duschinsky transformation given in equation 1.64. The development induced by this transformation is given in equation B.1. Using the generating function and replacing $\mathbf{Q}'$ by $\mathbf{Q}''$, the power series in equation 2.4 can be formulated:

$$
\begin{aligned}
f(\mathbf{T},\mathbf{U}) = \pi^{-N/2}\det\left[\boldsymbol{\Gamma}'\,\boldsymbol{\Gamma}''\right]^{1/4}\det(\mathbf{J})^{1/2}\exp\left[-\mathbf{U}^2-\mathbf{T}^2\right] \\
\times\int d\mathbf{Q}''\exp\Big[-\frac{1}{2}\big(\mathbf{Q}''^{\mathrm{T}}\mathbf{J}^{\mathrm{T}}\boldsymbol{\Gamma}'\mathbf{J}\mathbf{Q}''+\mathbf{Q}''^{\mathrm{T}}\mathbf{J}^{\mathrm{T}}\boldsymbol{\Gamma}'\mathbf{K}+\mathbf{K}^{\mathrm{T}}\boldsymbol{\Gamma}'\mathbf{J}\mathbf{Q}''+\mathbf{K}^{\mathrm{T}}\boldsymbol{\Gamma}'\mathbf{K} \\
+\mathbf{Q}''^{\mathrm{T}}\boldsymbol{\Gamma}''\mathbf{Q}''\big)+2\mathbf{U}^{\mathrm{T}}\boldsymbol{\Gamma}''^{1/2}\mathbf{Q}''+2\mathbf{T}^{\mathrm{T}}\boldsymbol{\Gamma}'^{1/2}\mathbf{J}\mathbf{Q}''+2\mathbf{T}^{\mathrm{T}}\boldsymbol{\Gamma}'^{1/2}\mathbf{K}\Big]
\end{aligned}
\tag{2.6}
$$

The same approach used in section 1.4 can be applied. However, because calculations are rather long, all the developments will be treated in appendix C and only the result will be given here.

For convenience, we report here the expression of the Sharp and Rosenstock matrices, $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, $\mathbf{D}$ and $\mathbf{E}$:

$$\mathbf{A} = 2\boldsymbol{\Gamma}'^{1/2}\mathbf{J}(\mathbf{J}^{\mathrm{T}}\boldsymbol{\Gamma}'\mathbf{J}+\boldsymbol{\Gamma}'')^{-1}\mathbf{J}^{\mathrm{T}}\boldsymbol{\Gamma}'^{1/2}-\mathbf{I} \tag{2.7}$$

$$\mathbf{B} = -2\boldsymbol{\Gamma}'^{1/2}(\mathbf{J}(\mathbf{J}^{\mathrm{T}}\boldsymbol{\Gamma}'\mathbf{J}+\boldsymbol{\Gamma}'')^{-1}\mathbf{J}^{\mathrm{T}}\boldsymbol{\Gamma}'-\mathbf{I})\mathbf{K} \tag{2.8}$$

$$\mathbf{C} = 2\boldsymbol{\Gamma}''^{1/2}(\mathbf{J}^{\mathrm{T}}\boldsymbol{\Gamma}'\mathbf{J}+\boldsymbol{\Gamma}'')^{-1}\boldsymbol{\Gamma}''^{1/2}-\mathbf{I} \tag{2.9}$$

$$\mathbf{D} = -2\boldsymbol{\Gamma}''^{1/2}(\mathbf{J}^{\mathrm{T}}\boldsymbol{\Gamma}'\mathbf{J}+\boldsymbol{\Gamma}'')^{-1}\mathbf{J}^{\mathrm{T}}\boldsymbol{\Gamma}''^{1/2}\mathbf{K} \tag{2.10}$$

$$\mathbf{E} = 4\boldsymbol{\Gamma}''^{1/2}(\mathbf{J}^{\mathrm{T}}\boldsymbol{\Gamma}'\mathbf{J}+\boldsymbol{\Gamma}'')^{-1}\mathbf{J}^{\mathrm{T}}\boldsymbol{\Gamma}'^{1/2} \tag{2.11}$$

Each formula to compute the overlap integrals is generated using the identity obtained from the two expressions of the power series given in equations 2.3 and C.13:

$$
\begin{aligned}
\sum_{\boldsymbol{v}'}\sum_{\boldsymbol{v}''}\mathbf{T}^{\boldsymbol{v}'}\mathbf{U}^{\boldsymbol{v}''}\left(\frac{2^{\boldsymbol{v}'}2^{\boldsymbol{v}''}}{\boldsymbol{v}'!\boldsymbol{v}''!}\right)^{1/2}\langle\,\boldsymbol{v}'\mid\boldsymbol{v}''\,\rangle \\
=\langle\,\mathbf{0}'\mid\mathbf{0}''\,\rangle\exp\Big[\mathbf{U}^{\mathrm{T}}\mathbf{C}\mathbf{U}+\mathbf{D}^{\mathrm{T}}\mathbf{U}+\mathbf{T}^{\mathrm{T}}\mathbf{A}\mathbf{T}+\mathbf{B}^{\mathrm{T}}\mathbf{T}+\mathbf{U}^{\mathrm{T}}\mathbf{E}\mathbf{T}\Big]
\end{aligned}
\tag{2.12}
$$

In this form, the equalization of the coefficients of powers of the dummy variables $\mathbf{U}$ and $\mathbf{T}$ to

formulate the overlap integrals is difficult. By switching back from matrix notation to a standard notation and replacing the exponential by the equivalent power series, the previous identity can be expressed in a more manageable layout:

$$\begin{aligned}
&\sum_{v'_1,\dots,v'_N=0}^{\infty}\sum_{v''_1,\dots,v''_N=0}^{\infty}\prod_{i=1}^{N}\mathrm{T}_i^{v'_i}\prod_{j=1}^{N}\mathrm{U}_j^{v''_j}\left(\prod_{i=1}^{N}\frac{2^{v'_i}2^{v''_i}}{v'_i!v''_i!}\right)^{1/2}\langle\,\boldsymbol{v'}\mid\boldsymbol{v''}\,\rangle \\
&\quad=\langle\,\mathbf{0'}\mid\mathbf{0''}\,\rangle\sum_{n=0}^{\infty}\frac{\left[\sum_{i=1}^{N}\mathrm{D}_i\mathrm{U}_i+\mathrm{B}_i\mathrm{T}_i+\sum_{j=1}^{N}\mathrm{U}_i\mathrm{C}_{ij}\mathrm{U}_j+\mathrm{T}_i\mathrm{A}_{ij}\mathrm{T}_j+\mathrm{U}_i\mathrm{E}_{ij}\mathrm{T}_j\right]^n}{n!}
\end{aligned} \tag{2.13}$$

In principle, equation 2.13 gives all possible combinations of overlaps at once. Practically, only one integral is calculated at a time. In this case, the sums in the left-hand side of the equations are unnecessary and can be removed, so that only one term remains there. As explained before, to find the formula of a given overlap integral $\langle\,\boldsymbol{v'}\mid\boldsymbol{v''}\,\rangle$, we need to look for all the coefficients of exactly the same combination of powers of T_i and U_j. Because of this condition, one can see that n does not need to be developed further than a limit equal to $\sum_{i=1}^{N}(v'_i+v''_i)$.

Finally, the sums over i and j in the right-hand side of equation 2.13 can be restricted to the value of the indexes corresponding to non-zero quantum numbers. It should be noted that the way we wrote the summation is not adapted to this process and to do so, one should write the sums separately for each term.

To clarify these explanations, let us consider a practical example where one wants to compute the overlap $\langle\,\mathbf{0'}\mid\mathbf{0''}+2''_i+1''_j\,\rangle$. This corresponds to the general case of a transition from the vibrational ground state of the initial state to a combinations of two modes with the quantum numbers : $v''_i=2, v''_j=1$.

In this case, all terms with the dummy variable $\mathbf{T}$ can be removed and equation 2.13 is the simple identity:

$$\begin{aligned}
&\mathrm{U}_i^2\mathrm{U}_j^1\left(\frac{2^2 2^1}{2!1!}\right)^{1/2}\langle\,\mathbf{0'}\mid\mathbf{0''}+2''_i+1''_j\,\rangle \\
&\quad=\langle\,\mathbf{0'}\mid\mathbf{0''}\,\rangle\sum_{n=0}^{3}\frac{\left[\mathrm{D}_i\mathrm{U}_i+\mathrm{D}_j\mathrm{U}_j+\mathrm{U}_i\mathrm{C}_{ii}\mathrm{U}_i+\mathrm{U}_i\mathrm{C}_{ij}\mathrm{U}_j+\mathrm{U}_j\mathrm{C}_{jj}\mathrm{U}_j+\mathrm{U}_j\mathrm{C}_{ji}\mathrm{U}_i\right]^n}{n!}
\end{aligned} \tag{2.14}$$

The terms in the right-hand side of equation 2.13 are simplified to only consider the relevant indexes i and j. After developing the polynomial in the right-hand side, the following identity is obtained:

$$4^{1/2}\times\langle\,\mathbf{0'}\mid\mathbf{0''}+2''_i+1''_j\,\rangle=\langle\,\mathbf{0'}\mid\mathbf{0''}\,\rangle\times\frac{1}{2}\times\left[{\mathrm{D}_i}^2\mathrm{D}_j+2\mathrm{C}_{ii}\mathrm{D}_j+2\mathrm{C}_{ij}\mathrm{D}_i+2\mathrm{C}_{ji}\mathrm{D}_i\right] \tag{2.15}$$

$\mathbf{C}$ is symmetrical so that $\mathrm{C}_{ji}=\mathrm{C}_{ij}$. Hence, the analytic formula of the overlap integral $\langle\,\mathbf{0'}\mid\mathbf{0''}+2''_i+1''_j\,\rangle$ is:

$$\langle\,\mathbf{0'}\mid\mathbf{0''}+2''_i+1''_j\,\rangle=\frac{1}{4}\times\langle\,\mathbf{0'}\mid\mathbf{0''}\,\rangle\times\left[{\mathrm{D}_i}^2\mathrm{D}_j+2\mathrm{C}_{ii}\mathrm{D}_j+4\mathrm{C}_{ij}\mathrm{D}_i\right] \tag{2.16}$$

Using this approach and generalizing it to each overlap integral is very simple in theory. However, it is very complex in practice to develop the right-hand side of equation 2.13 when the quantum numbers are high and the number of modes in the combination bands increase. As a consequence, the number of terms to compute the overlap integrals grows steeply. This induces risks of errors, be it by forgetting terms or mistyping them.

These errors can be avoided by using specialized computational tools for symbolic algebra [77]. They can be used to treat almost automatically the large equations resulting from taking into account combinations of many simultaneously excited modes with higher numbers of quanta.

2.4 A recursive method based on the Sharp-Rosenstock approach

The original method proposed by Sharp and Rosenstock is purely analytic [43]. However, recursion equations based on this model were devised to remove this potential drawback, a strong inadequacy to treat generic problems of medium to large systems.

We will present in this section a recursive approach proposed by Ruhoff [60] which uses a formalism similar to Sharp and Rosenstock's. A first change occurs in the expression of the generating function of the Hermite polynomial given in equation 2.5. The values of the dummy variables $\mathbf{U}$ and $\mathbf{T}$ being totally free, one can assume that they are near zero. Doing so, the exponential of the right-hand side of equation 2.5 can be expressed with a Maclaurin series, a particular case of the Taylor series. In this case, equation 2.5 can be written in the following form:

$$\sum_{\boldsymbol{v}''=0}^{\infty} \frac{\mathbf{U}^{\boldsymbol{v}''}}{\boldsymbol{v}''!} H_{\boldsymbol{v}''}(\boldsymbol{\Gamma}''^{1/2}\mathbf{Q}'') = \left(\sum_{\boldsymbol{v}''=0}^{\infty} \frac{\mathbf{U}^{\boldsymbol{v}''}}{\boldsymbol{v}''!} \frac{\partial^{\boldsymbol{v}''}}{\partial \mathbf{U}^{\boldsymbol{v}''}} \exp\left[-\mathbf{U}^2 + 2\mathbf{U}^{\mathrm{T}}\boldsymbol{\Gamma}''^{1/2}\mathbf{Q}''\right] \right)_{\mathbf{U}=\mathbf{0}} \tag{2.17}$$

$$\Rightarrow H_{\boldsymbol{v}''}(\boldsymbol{\Gamma}''^{1/2}\mathbf{Q}'') = \left(\frac{\partial^{\boldsymbol{v}''}}{\partial \mathbf{U}^{\boldsymbol{v}''}} \exp\left[-\mathbf{U}^2 + 2\mathbf{U}^{\mathrm{T}}\boldsymbol{\Gamma}''^{1/2}\mathbf{Q}''\right] \right)_{\mathbf{U}=\mathbf{0}} \tag{2.18}$$

In the same way as before, the convenient matrix notation is used at the expense of a correct mathematical formulation. The partial derivative given here corresponds in reality to a product of derivatives:

$$\frac{\partial^{\boldsymbol{v}''}}{\partial \mathbf{U}^{\boldsymbol{v}''}} = \frac{\partial^{v_1''}}{\partial \mathrm{U}_1^{v_1''}} \frac{\partial^{v_2''}}{\partial \mathrm{U}_2^{v_2''}} \cdots \frac{\partial^{v_N''}}{\partial \mathrm{U}_N^{v_N''}} \tag{2.19}$$

For the purposes of concision and clarity, we will continue to use the matrix notation, which does not alter the validity of our demonstration.

The generating function given in equation 2.18 for the Hermite polynomial, by setting some conditions on the dummy variables, does not require anymore the powers of $\mathbf{U}$ to be used with respect to the expression chosen by Sharp and Rosenstock in equation 2.5. A direct consequence of this is the possibility to express plainly the overlap integral $\langle\, \boldsymbol{v}' \mid \boldsymbol{v}'' \,\rangle$ without the need of a power series. Introducing equation 2.18 in the formula of the overlap integral given in equation 2.1, we

obtain the following relation, using the same general matrix notation as in equation 2.3:

$$\begin{aligned}\langle\, \boldsymbol{v}' \mid \boldsymbol{v}'' \,\rangle = \pi^{-N/2} \det\left[\boldsymbol{\Gamma}' \boldsymbol{\Gamma}''\right]^{1/4} \det(\mathbf{J})^{1/2} \left(2^{\boldsymbol{v}'} 2^{\boldsymbol{v}''} \boldsymbol{v}'! \boldsymbol{v}''!\right)^{-1/2} \int d\mathbf{Q}'' \\ \left(\frac{\partial^{\boldsymbol{v}'}}{\partial \mathbf{T}^{\boldsymbol{v}'}} \exp\left[-\mathbf{T}^2 + 2\mathbf{T}^{\mathrm{T}} \boldsymbol{\Gamma}'^{1/2} \mathbf{Q}'\right] \frac{\partial^{\boldsymbol{v}''}}{\partial \mathbf{U}^{\boldsymbol{v}''}} \exp\left[-\mathbf{U}^2 + 2\mathbf{U}^{\mathrm{T}} \boldsymbol{\Gamma}''^{1/2} \mathbf{Q}''\right] \right)_{\mathbf{T}=\mathbf{U}=\mathbf{0}} \\ \exp\left[-\frac{1}{2}(\mathbf{Q}'^{\mathrm{T}} \boldsymbol{\Gamma}' \mathbf{Q}' + \mathbf{Q}''^{\mathrm{T}} \boldsymbol{\Gamma}'' \mathbf{Q}'')\right]\end{aligned} \tag{2.20}$$

By reordering the terms in the integral after having used the Duschinsky transformation to replace $\mathbf{Q}'$, we can reach a more familiar expression:

$$\begin{aligned}\langle\, \boldsymbol{v}' \mid \boldsymbol{v}'' \,\rangle = \pi^{-N/2} \det\left[\boldsymbol{\Gamma}' \boldsymbol{\Gamma}''\right]^{1/4} \det(\mathbf{J})^{1/2} \left(2^{\boldsymbol{v}'} 2^{\boldsymbol{v}''} \boldsymbol{v}'! \boldsymbol{v}''!\right)^{-1/2} \\ \times \left(\frac{\partial^{\boldsymbol{v}'}}{\partial \mathbf{T}^{\boldsymbol{v}'}} \frac{\partial^{\boldsymbol{v}''}}{\partial \mathbf{U}^{\boldsymbol{v}''}} \exp\left[-\mathbf{T}^2 - \mathbf{U}^2\right] \int d\mathbf{Q}'' \exp\Big[\frac{1}{2}(\mathbf{Q}''^{\mathrm{T}} \mathbf{J}^{\mathrm{T}} \boldsymbol{\Gamma}' \mathbf{J} \mathbf{Q}'' \right. \\ + \mathbf{Q}''^{\mathrm{T}} \mathbf{J}^{\mathrm{T}} \boldsymbol{\Gamma}' \mathbf{K} + \mathbf{K}^{\mathrm{T}} \boldsymbol{\Gamma}' \mathbf{J} \mathbf{Q}'' + \mathbf{K}^{\mathrm{T}} \boldsymbol{\Gamma}' \mathbf{K} + \mathbf{Q}''^{\mathrm{T}} \boldsymbol{\Gamma}'' \mathbf{Q}'') \\ \left. + 2\mathbf{U}^{\mathrm{T}} \boldsymbol{\Gamma}''^{1/2} \mathbf{Q}'' + 2\mathbf{T}^{\mathrm{T}} \boldsymbol{\Gamma}'^{1/2} \mathbf{J} \mathbf{Q}'' + 2\mathbf{T}^{\mathrm{T}} \boldsymbol{\Gamma}'^{1/2} \mathbf{K}\Big] \right)_{\mathbf{T}=\mathbf{U}=\mathbf{0}}\end{aligned} \tag{2.21}$$

In this form, the equation is similar to the power series given in equation 2.6. By establishing this parallel to the method of Sharp and Rosenstock, we can directly use the results we obtained there, taking care of adapting them to our slightly different case.

As a consequence, the previous equation of the overlap integral can be simply integrated to the following form:

$$\begin{aligned}\langle\, \boldsymbol{v}' \mid \boldsymbol{v}'' \,\rangle = \left(2^{\boldsymbol{v}'} 2^{\boldsymbol{v}''} \boldsymbol{v}'! \boldsymbol{v}''!\right)^{-1/2} \left(\frac{\partial^{\boldsymbol{v}'}}{\partial \mathbf{T}^{\boldsymbol{v}'}} \frac{\partial^{\boldsymbol{v}''}}{\partial \mathbf{U}^{\boldsymbol{v}''}} \langle\, \mathbf{0}' \mid \mathbf{0}'' \,\rangle \right. \\ \left. \times \exp\left[\mathbf{U}^{\mathrm{T}} \mathbf{C} \mathbf{U} + \mathbf{D}^{\mathrm{T}} \mathbf{U} + \mathbf{T}^{\mathrm{T}} \mathbf{A} \mathbf{T} + \mathbf{B}^{\mathrm{T}} \mathbf{T} + \mathbf{U}^{\mathrm{T}} \mathbf{E} \mathbf{T}\right] \right)_{\mathbf{T}=\mathbf{U}=\mathbf{0}}\end{aligned} \tag{2.22}$$

This equation represents the starting point to generate our recursion formulae. The method we will employ here is derived from the work of Lermé [62] who presented a method to obtain recursion relations for mono- and bidimensional Franck-Condon integrals.

The exponential in the right-hand side of equation 2.22 can be differentiated with respect to U_i and T_i with i representing any mode ($i \in \{1, \ldots, N\}$). A complete differentiation is not necessary as we are looking for relations between overlap integrals. As a matter of fact, we will consider two independant kinds of partial differentiations, initiating the treatment by a derivative with respect to U_i, that will be studied first, and T_i, which will be the second case.

In the following equations, we will use a mixed notation, abandoning the matrix notation when necessary. In addition, the exponential stays untouched by differentiation so that it can be replaced by a function $F(\mathbf{T}, \mathbf{U}) = \exp\left[\mathbf{U}^{\mathrm{T}} \mathbf{C} \mathbf{U} + \mathbf{D}^{\mathrm{T}} \mathbf{U} + \mathbf{T}^{\mathrm{T}} \mathbf{A} \mathbf{T} + \mathbf{B}^{\mathrm{T}} \mathbf{T} + \mathbf{U}^{\mathrm{T}} \mathbf{E} \mathbf{T}\right]$. Because the derivatives of $F(\mathbf{T}, \mathbf{U})$ will be with respect to elements of $\mathbf{U}$ and $\mathbf{T}$, we will recall that $F(\mathbf{T}, \mathbf{U})$ is a relation

of the form:

$$F(\mathbf{T},\mathbf{U}) = \exp\Big[\sum_{i=1}^{N} \mathrm{D}_i\mathrm{U}_i + \mathrm{B}_i\mathrm{T}_i + \sum_{j=1}^{N} \mathrm{U}_i\mathrm{C}_{ij}\mathrm{U}_j + \mathrm{T}_i\mathrm{A}_{ij}\mathrm{T}_j + \mathrm{U}_i\mathrm{E}_{ij}\mathrm{T}_j\Big] \tag{2.23}$$

Let us differentiate $F(\mathbf{T},\mathbf{U})$ with respect to U_i so that equation 2.22 can be written:

$$\begin{aligned}\langle\, \boldsymbol{v}' \mid \boldsymbol{v}'' \,\rangle = \Big(2^{\boldsymbol{v}'}2^{\boldsymbol{v}''}\boldsymbol{v}'!\boldsymbol{v}''!\Big)^{-1/2}\Bigg(&\frac{\partial^{v'_1}}{\partial \mathrm{T}_1^{v'_1}}\cdots\frac{\partial^{v'_N}}{\partial \mathrm{T}_N^{v'_N}}\frac{\partial^{v''_1}}{\partial \mathrm{U}_1^{v''_1}}\cdots\frac{\partial^{v''_i-1}}{\partial \mathrm{U}_i^{v''_i-1}}\cdots\frac{\partial^{v''_N}}{\partial \mathrm{U}_N^{v''_N}}\langle\, \mathbf{0}' \mid \mathbf{0}'' \,\rangle \\ &\Big[\mathrm{D}_i + \big\{\sum_{j=1}^{N} 2\mathrm{C}_{ij}\mathrm{U}_j\big\} + \big\{\sum_{j=1}^{N} \mathrm{E}_{ij}\mathrm{T}_j\big\}\Big]F(\mathbf{T},\mathbf{U})\Bigg)_{\mathbf{T}=\mathbf{U}=\mathbf{0}}\end{aligned} \tag{2.24}$$

The right-hand side of equation 2.24 can be expanded into $2N+1$ terms, each one with a dependence on a dummy variable U_j or T_j, except for one. To eliminate this problem, $2N$ differentiations must be performed, one for each element U_j and T_j. To do so, we will use the Leibniz product rule which gives a general expression for arbitrary-order derivatives of products of functions:

$$\frac{d^n}{dx^n}f(x)g(x) = \sum_{k=0}^{n}\binom{n}{k}\left(\frac{d^i}{dx^i}f(x)\right)\left(\frac{d^{n-i}}{dx^{n-i}}g(x)\right) \tag{2.25}$$

where $\binom{n}{k}$ is the binomial coefficient.

In our present study, $f(x)$ would be "2 $\mathrm{C}_{ij}\mathrm{U}_j$" in the case a derivative with respect to U_j or "$\mathrm{E}_{ij}\mathrm{T}_j$" for T_j, and $g(x)$ would be $F(\mathbf{T},\mathbf{U})$. It is straightforward to see that the terms of the sum in equation 2.25 are null for $k \geq 2$. Given this fact, three relations can be derived from the Leibniz product rule:

- $\forall j \in \{1,\ldots,N\}$,
$$\frac{\partial^{v'_j}}{\partial \mathrm{T}_j^{v'_j}}\mathrm{E}_{ij}\mathrm{T}_j F(\mathbf{T},\mathbf{U}) = \mathrm{E}_{ij}\mathrm{T}_j\frac{\partial^{v'_j}}{\partial \mathrm{T}_j^{v'_j}}F(\mathbf{T},\mathbf{U}) + v'_j\mathrm{E}_{ij}\frac{\partial^{v'_j-1}}{\partial \mathrm{T}_j^{v'_j-1}}F(\mathbf{T},\mathbf{U}) \tag{2.26}$$
- $\forall j \in \{1,\ldots,i-1,i+1,\ldots,N\}$,
$$\frac{\partial^{v''_j}}{\partial \mathrm{U}_j^{v''_j}}\mathrm{C}_{ij}\mathrm{U}_j F(\mathbf{T},\mathbf{U}) = 2\mathrm{C}_{ij}\mathrm{U}_j\frac{\partial^{v''_j}}{\partial \mathrm{U}_j^{v''_j}}F(\mathbf{T},\mathbf{U}) + 2v''_j\mathrm{C}_{ij}\frac{\partial^{v''_j-1}}{\partial \mathrm{U}_j^{v''_j-1}}F(\mathbf{T},\mathbf{U}) \tag{2.27}$$
- $$\frac{\partial^{v''_i-1}}{\partial \mathrm{U}_i^{v''_i-1}}2\mathrm{C}_{ij}\mathrm{U}_j F(\mathbf{T},\mathbf{U}) = 2\mathrm{C}_{ii}\mathrm{U}_i\frac{\partial^{v''_i}}{\partial \mathrm{U}_i^{v''_i}}F(\mathbf{T},\mathbf{U}) + 2(v''_i-1)\mathrm{C}_{ii}\frac{\partial^{v''_i-2}}{\partial \mathrm{U}_i^{v''_i-2}}F(\mathbf{T},\mathbf{U}) \tag{2.28}$$

Using the three conditions above, the overlap integral can be written:

$$
\begin{aligned}
\langle\, \boldsymbol{v}' \mid \boldsymbol{v}'' \,\rangle = &\left(2^{\boldsymbol{v}'}2^{\boldsymbol{v}''}\boldsymbol{v}'!\boldsymbol{v}''!\right)^{-1/2}\langle\, \mathbf{0}' \mid \mathbf{0}'' \,\rangle \\
&\times \Bigg(\Big[\mathrm{D}_i + \{\sum_{j=1}^{N} 2\mathrm{C}_{ij}\mathrm{U}_j\} + \{\sum_{j=1}^{N} \mathrm{E}_{ij}\mathrm{T}_j\} \Big] \\
&\quad \times \frac{\partial^{v'_1}}{\partial \mathrm{T}_1^{v'_1}} \cdots \frac{\partial^{v'_N}}{\partial \mathrm{T}_N^{v'_N}} \frac{\partial^{v''_1}}{\partial \mathrm{U}_1^{v''_1}} \cdots \frac{\partial^{v''_i-1}}{\partial \mathrm{U}_i^{v''_i-1}} \cdots \frac{\partial^{v''_N}}{\partial \mathrm{U}_N^{v''_N}} F(\mathbf{T},\mathbf{U}) \\
&\quad + \sum_{j=1}^{N} v'\mathrm{E}_{ij} \frac{\partial^{v'_1}}{\partial \mathrm{T}_1^{v'_1}} \cdots \frac{\partial^{v'_j-1}}{\partial \mathrm{T}_j^{v'_j-1}} \cdots \frac{\partial^{v'_N}}{\partial \mathrm{T}_N^{v'_N}} \frac{\partial^{v''_1}}{\partial \mathrm{U}_1^{v''_1}} \cdots \frac{\partial^{v''_i-1}}{\partial \mathrm{U}_i^{v''_i-1}} \cdots \frac{\partial^{v''_N}}{\partial \mathrm{U}_N^{v''_N}} F(\mathbf{T},\mathbf{U}) \\
&\quad + \sum_{\substack{j=1 \\ j\neq i}}^{N} 2v'_j\mathrm{C}_{ij} \frac{\partial^{v'_1}}{\partial \mathrm{T}_1^{v'_1}} \cdots \frac{\partial^{v'_N}}{\partial \mathrm{T}_N^{v'_N}} \frac{\partial^{v''_1}}{\partial \mathrm{U}_1^{v''_1}} \cdots \frac{\partial^{v''_i-1}}{\partial \mathrm{U}_i^{v''_i-1}} \cdots \frac{\partial^{v''_j-1}}{\partial \mathrm{U}_j^{v''_j-1}} \cdots \frac{\partial^{v''_N}}{\partial \mathrm{U}_N^{v''_N}} F(\mathbf{T},\mathbf{U}) \\
&\quad + 2(v'_i-1)\mathrm{C}_{ii} \frac{\partial^{v'_1}}{\partial \mathrm{T}_1^{v'_1}} \cdots \frac{\partial^{v'_N}}{\partial \mathrm{T}_N^{v'_N}} \frac{\partial^{v''_1}}{\partial \mathrm{U}_1^{v''_1}} \cdots \frac{\partial^{v''_i-2}}{\partial \mathrm{U}_i^{v''_i-2}} \cdots \frac{\partial^{v''_N}}{\partial \mathrm{U}_N^{v''_N}} F(\mathbf{T},\mathbf{U}) \Bigg)_{\mathbf{T}=\mathbf{U}=\mathbf{0}}
\end{aligned}
\tag{2.29}
$$

By applying the conditions on the dummy variables ($\mathbf{U} = \mathbf{T} = \mathbf{0}$), we can simplify the first term of the right-hand side of equation 2.29 and only keep the factor D_i. Analyzing all the terms in the right-hand side, we can see that each one is very similar to the formula of a Franck-Condon integral given in equation 2.22, the only needed correction being the factor $(2^{\boldsymbol{v}'}2^{\boldsymbol{v}''}\boldsymbol{v}'!\boldsymbol{v}''!)^{-1/2}$. To illustrate this, let us consider a single term of the previous equation and transform it into a Franck-Condon integral:

$$
\begin{aligned}
&\left(2^{\boldsymbol{v}'}2^{\boldsymbol{v}''}\boldsymbol{v}'!\boldsymbol{v}''!\right)^{-1/2}\langle\, \mathbf{0}' \mid \mathbf{0}'' \,\rangle \times 2 \times v'_j \times \mathrm{C}_{ij} \\
&\qquad \times \frac{\partial^{v'_1}}{\partial \mathrm{T}_1^{v'_1}} \cdots \frac{\partial^{v'_N}}{\partial \mathrm{T}_N^{v'_N}} \frac{\partial^{v''_1}}{\partial \mathrm{U}_1^{v''_1}} \cdots \frac{\partial^{v''_i-1}}{\partial \mathrm{U}_i^{v''_i-1}} \cdots \frac{\partial^{v''_j-1}}{\partial \mathrm{U}_j^{v''_j-1}} \cdots \frac{\partial^{v''_N}}{\partial \mathrm{U}_N^{v''_N}} F(\mathbf{T},\mathbf{U}) \\
&= 2v'_j\mathrm{C}_{ij}\left(2v''_i2v''_j\right)^{-1/2}\Big(2^{v'_1}\ldots 2^{v'_N}2^{v''_1}\ldots 2^{v''_i-1}\ldots 2^{v''_j-1}\ldots 2^{v''_N} \\
&\qquad v'_1!\ldots v'_N!v''_1!\ldots(v''_i-1)!\ldots(v''_j-1)!\ldots v''_N!\Big)^{-1/2}\langle\, \mathbf{0}' \mid \mathbf{0}'' \,\rangle \\
&\qquad \frac{\partial^{v'_1}}{\partial \mathrm{T}_1^{v'_1}} \cdots \frac{\partial^{v'_N}}{\partial \mathrm{T}_N^{v'_N}} \frac{\partial^{v''_1}}{\partial \mathrm{U}_1^{v''_1}} \cdots \frac{\partial^{v''_i-1}}{\partial \mathrm{U}_i^{v''_i-1}} \cdots \frac{\partial^{v''_j-1}}{\partial \mathrm{U}_j^{v''_j-1}} \cdots \frac{\partial^{v''_N}}{\partial \mathrm{U}_N^{v''_N}} F(\mathbf{T},\mathbf{U}) \\
&= \left(\frac{v''_j}{v''_i}\right)^{1/2} \mathrm{C}_{ij}\langle\, \boldsymbol{v}' \mid \boldsymbol{v}'' - 1''_i - 1''_j \,\rangle
\end{aligned}
\tag{2.30}
$$

The procedure presented above can be systematized for each term of the right-hand side of equation 2.29. By doing so, we can obtain a simpler form which will be one of our two recursion formulae:

$$
\begin{aligned}
\langle\, \boldsymbol{v}' \mid \boldsymbol{v}'' \,\rangle = \frac{1}{\sqrt{2v''_i}}\Big[&\mathrm{D}_i\langle\, \boldsymbol{v}' \mid \boldsymbol{v}'' - 1''_i \,\rangle + \sqrt{2(v''_i-1)}\,\mathrm{C}_{ii}\langle\, \boldsymbol{v}' \mid \boldsymbol{v}'' - 2''_i \,\rangle \\
&+ \sum_{\substack{j=1 \\ j\neq i}}^{N} \sqrt{2v''_j}\,\mathrm{C}_{ij}\langle\, \boldsymbol{v}' \mid \boldsymbol{v}'' - 1''_i - 1''_j \,\rangle + \sum_{j=1}^{N} \sqrt{v'_j}\,\mathrm{E}_{ij}\langle\, \boldsymbol{v}' - 1'_j \mid \boldsymbol{v}'' - 1''_i \,\rangle\Big]
\end{aligned}
\tag{2.31}
$$

where $(2v_i'')^{-1/2}$ is a common factor from the initial differentiation.

In the same way, starting by differentiating equation 2.22 with respect to T_i, we can generate a second, complementary recursion equation. The development being equivalent to what has been presented here, only the result will be given:

$$\begin{aligned}\langle\, \boldsymbol{v}' \mid \boldsymbol{v}'' \,\rangle = \frac{1}{\sqrt{2v_i'}}\Big[&\mathrm{B}_i\langle\, \boldsymbol{v}' - 1_i' \mid \boldsymbol{v}'' \,\rangle + \sqrt{2(v_i'-1)}\;\mathrm{A}_{ii}\langle\, \boldsymbol{v}' - 2_i' \mid \boldsymbol{v}'' \,\rangle \\ &+ \sum_{\substack{j=1\\ j\neq i}}^{N} \sqrt{2v_j'}\;\mathrm{A}_{ij}\langle\, \boldsymbol{v}' - 1_i' - 1_j' \mid \boldsymbol{v}'' \,\rangle + \sum_{j=1}^{N} \sqrt{v_j'}\;\mathrm{E}_{ji}\langle\, \boldsymbol{v}' - 1_i' \mid \boldsymbol{v}'' - 1_j'' \,\rangle\Big]\end{aligned} \tag{2.32}$$

In the general case, several vibrational states are populated in the initial state ($T > 0$ K) and both equations 2.31 and 2.32 are necessary. The second one is more general and allows to express an overlap integral with respect to overlap integrals involving lower quanta for the initial state. The first relation is generally used when the transition starts from the fundamental vibrational state of the initial state.

Practically, to express a given overlap integral $\langle\, \boldsymbol{v}' \mid \boldsymbol{v}'' \,\rangle$ with non null $\boldsymbol{v}'$ and $\boldsymbol{v}''$, according to the overlap integral $\langle\, \mathbf{0}' \mid \mathbf{0}'' \,\rangle$ the recursion is done in two main steps. Firstly, equation 2.32 is used until the fundamental vibrational state of the initial state ($\boldsymbol{v}' = \mathbf{0}'$) is "reached" and finally equation 2.31 to express every resulting overlap integral with respect to $\langle\, \mathbf{0}' \mid \mathbf{0}'' \,\rangle$.

In the absence of excitation provoked by the temperature, only the fundamental vibrational state of the initial state is populated and equation 2.31, without the last term in the right-hand side, is enough to perform all recursive calculations.

2.5 Aiming at an efficient calculation

From the different relations given previously, it is possible to compute any overlap integral. In the case of the analytic approach from Sharp and Rosenstock [43], and in general all methods of this kind, the expressions that need to be generated and calculated become rather complex for transitions between combinations bands. A direct consequence is that many implementations are often tailored to particular systems or groups of systems. A general-purpose method is more easily obtained and developed from recursion formulae. This implementation, however, is not without constraints due to the very principle of recursivity.

Some problems have been discussed briefly in section 2.2 but in this section, we will focus on two major difficulties which have been regularly discussed in works dealing with the generation of vibronically resolved spectra using the Franck-Condon principle. Because of the redundant calculations involved by the very principle of recursivity, it is often more efficient to store overlap integrals than to recompute them each time they are needed. Therefore, in a first part, we will discuss about the methods available to store these integrals and index them in memory. While this problem is mostly computer-centered, it has a major impact on the efficiency of programs based on recursion approaches and must be carefully treated to take full advantage of the computers performances. The second point is the limitation of the transitions to deal with. Theoretically, there is no limit to the value of the quantum number of each mode. This means that an infinity of overlap integrals should be taken into account, which is impossible to carry out. A procedure

for limiting the transitions to take into account is necessary and possible schemes to do so will be discussed afterwards. For each of these two issues, we will make a brief overview of solutions.

2.6 Indexing and recovering overlap integrals

The facility to evaluate each overlap integral given by recursivity is impeded by the constant re-calculations involved. As mentioned previously, the storage of the Franck-Condon integrals once they have been calculated can be a solution to this problem. However, a direct saving in memory is generally not a good solution. Indeed, considering equations 2.31 and 2.32, one can see that several kinds of integrals are needed and there is no certainty about the moment the latter were computed. Presented in another way, a straightforward storage of each integral can cause its recovery to be difficult when it has to be recalled. From this observation, two constraints can be formulated to optimize the efficiency of the calculations.

The first one, indirectly related to the storage, is that the order of the calculations is important in recurrence expressions. To understand this requirement, the simplest way is to consider the complete recursion to express a given overlap integral with respect to the overlap integral between the fundamental ground states. Let us consider the simple case of a non-linear triatomic molecule ($N = 3$) and a particular transition $\langle\, 1,0,0 \mid 2,1,0 \,\rangle$. This example is chosen to be as generic as possible and to make use of both recursion formulae. As explained in section 2.4, the more general relation given in equation 2.32 is applied first and then the second relation from equation 2.32. For the latter, i is chosen to be the highest mode in energy with a non-zero quantum number. A simple diagram showing all necessary overlap integrals is given in figure 2.1.

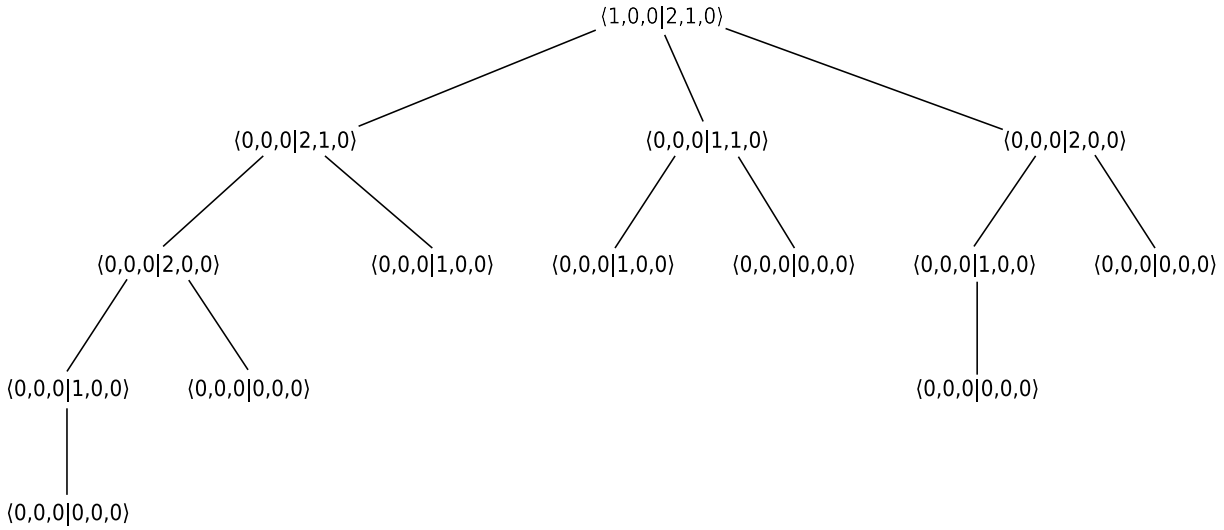


Figure 2.1: Recursion scheme for the transition $\langle\, 1,0,0 \mid 2,1,0 \,\rangle$ in the case of a non-linear triatomic molecule.

As shown in this figure, a complete recursive calculation will require several times the same overlap integrals (in this case, twice $\langle\, 0,0,0 \mid 2,0,0 \,\rangle$ and four times $\langle\, 0,0,0 \mid 1,0,0 \,\rangle$). To avoid this, storage is highly interesting, especially when the number of normal modes, and so the possible combination bands, increase. To be truly efficient, most, and preferably all, necessary overlap integrals should be available in memory, which requires that they have been computed before. To satisfy

this condition and also to avoid incessant lookups in the memory, one must take care of the order of calculations of the Franck-Condon integrals. From the recursion relations, the main restriction is to choose a consistent method to consider all necessary integrals. Several approaches are possible and are chosen to be coherent with the choice of the first differentiation. A straightforward one, used in our case, is to gradually increase the quantum number starting from the mode of lowest energy. Starting from mode 1, the quantum number is increased until a limit $v_{1_{\max}}$. When this limit is reached, v_2 is increased by one and v_1 reset to zero. The first quantum number is increased until it reaches the limit another time and the second quantum number is incremented by one. If it reaches the limit $v_{2_{\max}}$, then the third quantum number is set to one and so on. Each time, the mode corresponding to the first differentiation, i, is chosen to be the highest mode with a non-zero quantum number.

Having set a calculation scheme, a second constraint needs to be overcome, that is to say, finding a general way to store and recover previously calculated overlap integrals. Without going too deeply into the technical details which would be outside the scope of this work, it is important to note that several general modes exist to deal with recursive calculations. Berger *et al.* [68] proposed a classification in three methods, the last one having two variations. Historically, since the computer performances were limited, emphasis was put in a global storage of the data and so complex storage mecanisms were devised to save and index a possibly large number of integrals. Methods of this kind were referred as *conventional methods* in reference [68]. Because of the complexity of the most efficient indexing algorithms, their implementation can be difficult and require powerful mathematical and computer tools. Another approach favors a partial storage (*semidirect method*). In this case, only some integrals are saved and the rest is calculated if necessary. The burden on the memory is partly lifted with respect to an overall storage. Moreover, the integrals that need to be saved may be chosen freely and *ad hoc* methods can be devised, so that indexing is less cumbersome. Consequently, the efficiency of the data extraction/insertion is less critical for the performance of a procedure of simulation of Franck-Condon spectra. Finally, putting aside the storage and performing each calculation of overlap integrals with a complete reccurence expression corresponds to the *direct method*. In this case, calculations can be done starting from $\langle\, \boldsymbol{v'} \mid \boldsymbol{v''} \,\rangle$ to reach $\langle\, \mathbf{0'} \mid \mathbf{0''} \,\rangle$ as shown in figure 2.1 or the other way, starting from the transition between vibrational ground states. It is noteworthy that reducing the storage needs will increase the importance of the calculation order to avoid as much as possible redundant calculations and keep an optimum efficiency. Saving less overlap integrals also means that more calculations are required and so, as mentioned in section 2.2, a balance between calculations and storage needs to be found. It is noteworthy that the performances of the processors increase faster than the memory latency. As a consequence, the direct and semidirect methods are increasingly interesting. However, a pure, *direct method* is currently not easily efficient for large systems because of the huge redundance of calculations. As a result, *semidirect* or *conventional* methods are still preferable for a general-purpose program. These methods should be seen as general principles of working out the problem of the numerous computations involved by recursivity. To be practically effective, the choice of a *conventional* or *semidirect* kind of storage must be accompanied by a strategy to manage the memory and the saving of the transition intensities. Consequently, we will present some quite successful algorithms to index overlap integrals in memory. While these algorithms were originally devised for a global

storage, that is to say for a *conventional method*, they are also valid in the case of a partial storage. Taking into account this characteristic, we will firstly present the general principle of these indexing algorithms without regard to the chosen kind of *method*. However, the latter may have a noticeable influcence on their implementation inside the program, so we will indicate when it is the case.

The problem of indexing really started with the study of medium-size systems. Gruner and Brumer [67] considered the general case of polyatomic systems and proposed to store overlap integrals in a binary tree. Extensive details on the principles underlying the tree structure in computer science can be found in references [78, 79]. This structure has a strong analogy to actual trees and consists in a "branching" relation between nodes which represent addresses in the memory. A particular node corresponds to the starting point of the tree and is called the *root* of the tree. The nodes connected to it are referred as its children and each one is the start of a subset, also called *subtree*, disjoint from the others. The set of nodes being finite, some nodes do not have any children. These nodes are called *terminal nodes* or, still by analogy to the natural tree, *leaves*. In the particular case of the binary tree, each node can have at most two children, which can be conveniently referred as left subtree and right subtree.

Gruner and Brumer [67] proposed a simple and efficient method using binary trees to find any overlap by associating a change in mode i to the left subtrees and a change of the quantum number v_i to the right subtree. As an example, let us reconsider our triatomic molecule. However, this time we will assume that all transitions start from the vibrational ground state of the initial state for simplification. To generate our binary tree, let us fix the following limit: $v_{1_{\max}} = 2$, $v_{2_{\max}} = 1$, $v_{3_{\max}} = 2$ which will define the complete structure of our tree. The corresponding binary tree is represented in figure 2.2.

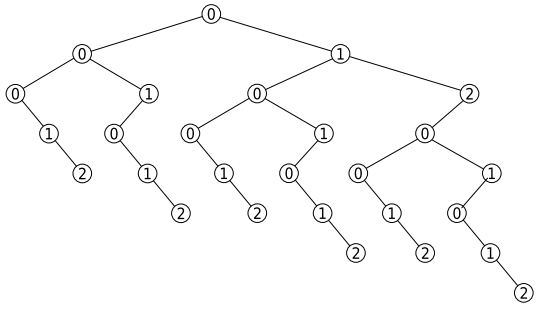


Figure 2.2: Example of a binary tree for a triatomic molecule. Only transitions from the vibrational ground states are considered here and maximum quantum numbers have been set to (2,1,2). The circled numbers correspond to the value of v_i.

In each node are stored the addresses of the left and right subtrees. In case there does not exist a left or right subtree, the corresponding address is null. Because of the structure of the tree, it is unnecessary to store the value of the quantum numbers (represented by the circled numbers in the figure 2.2). The procedure to find a specific overlap integral is as follows. Starting from the first mode at the root of the tree, the right branches are followed v_1 times. From the reached node, the left branch is followed once and then the right branches v_2 times. This last step, left once and right v_i times, is performed again until the last mode (N) is done. From this procedure, one can see that when the searched overlap is found, the corresponding node is not necessarily a *leaf* from the strict observance of its definition. The terminal nodes in our case will be all those without a left branch. In their original implementation, Gruner and Brumer [67] used two arrays. One was storing the overlap integrals and the other one, the binary tree, stored the index of the first array to find each overlap integral. In their terminal nodes, they stored these indexes in place of the left subtree address. In following implementations of binary trees [61, 69, 70], the first

array was discarded and the overlap integrals were directly stored in the "terminal nodes" of the binary tree.

The binary tree is not without strong drawbacks. The main problem is the high memory need. Gruner and Bruner proposed to solve it by using disk memory, which is unfortunately really slow to access. Ruhoff and Ratner [61] proposed to use the definition of the recursion formulae to restrict the binary tree to a chosen subset of overlap integrals. To classify the overlap integrals, they introduced the idea of levels "L" as the sum of all quantum numbers of the initial and final vibrational states. From this definition, the overlap integral of the group L can be calculated from the overlap integrals of the levels "L-1" and "L-2". Hence, overlap integrals from lower levels can be discarded, which reduces the burden on the memory. However, one can see that to be truly efficient, all integrals of a given level "L" should be calculated at the same time. This imposes new conditions on the order of calculation of the overlap integrals. Moreover, the decrease in memory usage is only partial and the memory requirement still grows steeply with the number of normal modes. Later, Hazra and Nooijen [69] used the approach presented by Ruhoff and Ratner [61] and proposed a way to improve it by discarding in each level the overlap integrals corresponding to the least probable transitions. While this reduced again the size of the binary tree, the memory usage could still exceed the resources available for the program so a restriction was introduced. If the number of integrals in a given level cannot be stored in the memory of a computer, the level is "divided" in sets of size smaller than the set limit and calculations are run for each set. This method has several disadvantages when applied to medium-to-large systems. Firstly, the threshold on the transition probabilities is arbitrarily set. As Hazra and Nooijen pointed out, calculations may need to be restarted from the beginning if the threshold is not set correctly and is too high and too many overlap integrals are discarded. Also, the limitations on the size of the levels can cause redundant calculations, some overlap integrals being recalculated several times when switching to another "set". While Hazra and Nooijen considered that the limitation on the size was a minor problem and should not be reached too much by each level, one could point out that this statement is strongly dependent on the available memory and on the number of normal modes. Dierksen and Grimme [70] used a similar procedure as Hazra and Nooijen but discarded the arbitrary threshold on the transition probability for the storage in the binary tree. Because all overlap integrals are stored, the risk of recalculation when the binary tree is "flushed" is very low. However, the memory usage is also higher and the resets can be more frequent if the available resources are limited.

The binary tree is a powerful approach for a *conventional method* to compute the overlap integrals but its memory consumption makes it tricky to use when the latter is limited. When implemented in a standalone program, the memory usage can be freely set within the limit of the available memory of the computer. However, when integration within another software is retained (as in our case), the memory consumption must be controlled to be within the limits fixed by the main program. As a result, such an approach is not well adapted to our needs.

Another approach, presented by Schumm *et al.* [80] relies on hash tables. The underlying theory is slightly less obvious to grasp than the trees previously discussed and we will only overview its general principle. More in-depth discussions about hashing can be found in references [78, 79]. In the case of the hash table, the storage structure is a common array. The particularity lies in the use of a transformation function, called *hash* function, which maps a set of keys on the array indexes.

The complexity of the method comes from the fact that in theory several keys can lead to the same array address. Hence, recovering the index from the key is not enough, one also needs to control that the index is the one looked for. When several keys lead to the same array address, this is referred as a collision and the risk of their occurrences need to be reduced. This is done through verification procedures whose complexity increases with the probability of collisions. The principal condition for an efficient hashing is that the hash function must, as much as possible, evenly distribute the keys on the array adresses, in particular to avoid the concentration of collisions on some zone of the memory.

To avoid too many collisions, which would imply complex collision resolutions methods, Schumm *et al.* [80] proposed to oversize the storage array with respect to the number of integrals to store. To create their hash functions, they used a number of approximations. The first and more important one is to set a constant maximum number of quanta for each mode, defined as "B-1". The hashing is then done in two steps. Firstly, a number of N figures in base "B" is created by placing side by side each quantum number defining the final state. It should be noted that in their work, Schumm *et al.* assumed that all transitions stemmed from the same vibrational state, which could however be different from the ground state. Additionally, they observed that the modes with lower energies were more likely to have a high number of quanta than modes of higher energies. Hence, the quantum numbers are ordered by decreasing energy from left to right. In a second step, the obtained number is converted in decimal base. This conversion explains their choice to "invert" the order of the modes so that the final result is as small as possible. As an example, we will use their procedure to retrieve the array index for a given final state $|\,2,1,0\,\rangle$. We will assume that the maximum number of quanta for each mode is 3, so B = 4. The complete sequence of hashing can be described with the following scheme:

$$|\,2,1,0\,\rangle \quad \rightarrow \quad (012)_4 \quad \rightarrow \quad 2_4 \times {4_{10}}^0 + 1_4 \times {4_{10}}^1 + 0_4 \times {4_{10}}^2 = 6_{10}$$

where the subscripts 4 and 10 represent the numeral base (so 10 corresponds to the decimal base).

It is straightforward to see that the obtained index can become huge. As a consequence, two problems can occur. The first one is that if one wishes to avoid collisions, an array whose size is at least "B^N" is needed. As Schumm *et al.* observed, this is practically unviable for medium-to-large systems, even on modern computer and collisions are unavoidable for a general treatment. Hence, a collision treatment is necessary to find the correct array address when such a case happens. A second problem is due to the limitations of the machine precision. An integer is generally limited to a maximum that it should never exceed[c]. For medium-size system, this limit can be easily reached, especially if "B" is chosen relatively high. To avoid this issue, Schumm *et al.* proposed some further approximations and a slightly more complex key. Considering the case of phenol, they proposed to keep the first six digits of the number obtained by placing side by side the quantum numbers and then discard all quantum numbers equal to zero. This method is efficient to reduce the value of the index but the evenness of the distribution is less controlled. Moreover, the risk of collisions is increased as several combinations will give exactly the same index.

[c]The behavior of a program when dealing with a number which overcome the machine precision depends on the programming language and is not always standardized. In any case, the program cannot keep the exact value of the number and the real index is then lost.

Because of the need of a transformation function and a collision control, the hash table seems less interesting than the trees structures. However, chosen correctly, the hash function can be really fast and tuned to specific needs, for example a limited available storage space. In the same way as binary trees and their flushing, the collision handling must be limited or, if possible, made unnecessary. It is rather interesting in the case of a *semidirect method* of recurrence calculations in which the set of overlap integrals to store is limited and has been chosen beforehand. Then, the hash function can be tailored to the conditions of application.

These interesting features and the relative freedom on the storage structure and size make the approaches based on a hash function very interesting for our included procedure. Taking into account the progress of processor performances and the possibility of parallel calculations, we can divide the overlap integrals to treat in *sets*. A *set* will be composed of all overlap integrals meeting two conditions on the final state. The first one is that the number of excited modes, that is to say the number of non-zero quantum numbers of the final state must be the same. The second one is that the excited modes must be the same. A *set* is defined by the maximum allowed quantum numbers reachable by each mode. Contrary to Schumm *et al.*, we do not impose that this maximum number of quanta must be the same for each mode.

As an example, let us consider a non-linear molecule with 4 atoms. Its number of vibrational modes is $N = 6$. For now, we will not consider the initial state and only treat the final state. We choose a particular set defined by the ket $|\,3,2,0,2,0,0\rangle$. From condition 1, an overlap integral with the final state given by $|\,2,1,0,1,0,0\,\rangle$ belongs to this set but not an overlap integral correponding to the transition to the state $|\,1,1,1,2,0,0\,\rangle$. The second condition will not be met if the final state is $|\,2,1,0,0,1,0\,\rangle$. This primary division of the complete set of overlap integrals makes possible some simplifications. Since each integral has the same number n of excited modes, we can create a first mapping of the index:

$$|\,2,1,0,1,0,0\,\rangle \quad \rightarrow \quad |\,2,1,1\,\rangle$$

In the perspective of an independent treatment of each set, this operation can be done before the actual calculations of the line intensity of each transition. The correspondance between the position of a given mode in the "shortened" list and its correct index in the vector $\boldsymbol{v}''$ is stored in a n-dimension list. Using the reduced lists of quantum numbers v_i'' and maximum quantum numbers $v''_{i_{\max}}$, the mapping function we used can be defined as the expression:

$$F(\underline{\boldsymbol{v}}'', \underline{\boldsymbol{v}}''_{\max}) = \sum_{i=1}^{n} \underline{v}_i'' \prod_{j=1}^{i-1} (\underline{v}''_{j\max} + 1) \tag{2.33}$$

where $\underline{\boldsymbol{v}}$ represents the subset obtained previously by discarding all zero quantum numbers. The function, through the term "$\underline{v}''_{j\max} + 1$", accounts for the storage of subsets corresponding to lower classes with the same excited modes as those of the reduced list of quantum number. This will be discussed in details in chapter 3 and more precisely in section 3.7.

In practice, and as we will show in chapter 4, n does not need to be high, a maximum of 10 excited modes being sufficient even for large systems to simulate accurately the transition spectra. Moreover, except some particular cases, $v''_{i_{\max}}$ is often small so that there is a relatively low probability that the generated index might be too high for the programming language upper limit

for integers. If needed, additional protections can be devised. The creation of small sets is also done to avoid the risk of collisions by allowing a unique storage address for each overlap integral. In this frame, our mapping function corresponds to a minimal perfect hash function and collision resolution is unnecessary.

In our discussion, we have neglected the initial state until now. The principle of our hashing procedure is to obtain a single index since our storage array is linear. The initial state can be simply treated by using a bi-dimensional array, each dimension representing one electronic state. The hash function given in equation 2.33 is used alternatively on the initial state and the final state depending on the needs. The memory usage becomes slightly more complex to manage but can be handled by increasing the memory available for the procedure or, if many initial vibrational states, include the possibility of collisions in the array.

The division in sets could also be used in a binary tree. However, the construction of the structure imposes a higher memory consumption with no important benefit with respect to the hash table.

2.7 Prescreening of the overlap integrals

While the issue of the storage has been resolved, calculations cannot be performed if we need to treat an infinity of transitions. Previously, we already assumed that we could determine a limit for each mode for the maximum number of quanta ($v_{i_{\max}}$) to take into account, in order to manage the indexing of the overlap integrals. It is indeed obvious that a method needs to be devised to avoid calculating non relevant transitions. An important characteristic of such a scheme is that it has to be consistent and not restricted to particular cases. We can define two kinds of approaches to limit calculations, those performed *a priori* [19, 70, 81–83], relying on prescreening through evaluation, and those done in real-time [67–69, 84]. By "real-time", we refer to the fact that the criterion is applied during the calculations of the transition probabilities. If this criterion is not met, calculations of some of the following integral will not be performed then.

Because they are the most obvious to consider and are in general easier to devise, we will firstly present some representative methods applied during the calculations. Maybe the most simple one uses the bounds of the spectrum as the energy criterion to determine which transitions need to be computed. More precisely, all combinations of quantum numbers of the initial and final states that correspond to energy differences falling into an interval defined by the lower and upper bound of the spectrum are evaluated. To be exact, the previous definition supposes that these bounds were relative to the energy of the transition between the vibrational ground states of the electronic states. To find all integrals satisfying these conditions, a consistent and thorough analysis must be carried out. Kemper et al. [85] proposed a backtracking algorithm to count all possible states for a giving energy interval. Contrary to most previous algorithms such as the one presented by Beyer and Swinehart [86], their method retained the information on the levels involved in the transition, rendering possible the calculation of the overlap integrals. While the procedure was designed for an arbitrary precision of the energy levels, including anharmonicity, we will present it here in the case of the harmonic approximation. In this context, the energies of the overtones are proportional to the number of quanta as written in equation 1.39. The counting starts from the first mode

which is supposed to be the lowest in energy[d], after all quantum numbers have been initialized to 0. The associated quantum number v_1'' is incremented and two controls are performed each time. If the corresponding energy is included in the interval of interest, then the transition to this final state can be treated. If the energy exceeds the upper limit, then the quantum number is decreased ($v_1'' = v_1'' - 1$) and the previous procedure is now applied to the second mode ($i = 2$). The sequence is repeated until the last mode, $i = N$, is treated. At this point, the algorithm looks for the highest integer j lower than N whose corresponding quantum number is non-zero ($v_j'' \neq 0$). The latter is decreased by one ($v_j'' = v_j'' - 1$) and the previous loop incrementing the quantum numbers is done on the interval $i = [j + 1, N]$. The complete procedure is run until $j = 0$.

Berger and Klessinger [84] pointed out that the original procedure did not take into account the initial state and so assumed that all transitions were starting from the vibrational ground state. To remedy to this problem, they proposed to use two backtracking procedures, one for each state. They named it an interlocked algorithm because one procedure, for the final states, run inside the other one treating the initial states. For each combination of quantum numbers of initial state corresponding to an energy inferior to the upper bound, the backtracking procedure described above is run on the combinations of final state. This is done until the backtracking procedures have reached their end in both states.

Because it is simple to implement and intuitive physically, this method, or similar counting methods, has been commonly used in many algorithms to compute Franck-Condon integrals and simulate absorption or emission spectra [67, 68, 70, 81]. Its main problem is the rapid drop in performance if a large interval of energies is requested for the spectum simulation. As a matter of fact, from the way the backtracking is constructed, the interval has a not a direct impact on the performance since the most important value is the upper bound of the spectrum for absorption, or the lower bound for emission. Consequently, for a case corresponding to a small interval near the energy of the transitions between the vibrational ground states, the performance can be very good. For a more general need, and if a large spectrum is required, the computational times increase steeply to reach undesirable heights as shown in figure 2.3.

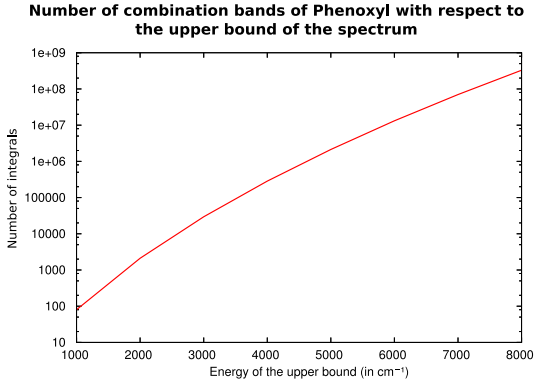


Figure 2.3: Case of the $^2B_1 \rightarrow {}^2A_2$ transition of the Phenoxyl radical. The number of transitions whose corresponding energies fall in the interval 0-upper_bound is reported with respect to the upper_bound energy. All transitions start from the vibrational ground state. Calculations were done at the TD-DFT level for the excited state (B3LYP/TZVP).

Hazra and Nooijen [69] proposed a different approach where the limitation is not set on the energy of the transition but on the analysis of the probability of transition. As we mentioned before when considering the problem of storage, Hazra and Nooijen used a threshold to discard the less likely transitions. By doing this, they also removed gradually the transitions to combinations bands

[d]The counting can also be carry out in a descending order; the main constraint is to keep the same order throughout the procedure.

of the same excited modes but with higher quanta. This behavior is a "natural" consequence of their treatment of the overlap integrals by increasing the level, that is to say the sum of all quantum numbers. This is done through a backward study of the recursion formulae given in equations 2.31 and 2.32. Instead of finding which overlap integrals are needed to calculate $\langle \boldsymbol{v}' \mid \boldsymbol{v}'' \rangle$, they looked for all integrals that could be calculated knowing $\langle \boldsymbol{v}' \mid \boldsymbol{v}'' \rangle$. Hence, starting from the level L_0 ($\langle \mathbf{0}' \mid \mathbf{0}' \rangle$), they listed all possible integrals from the level L_1 which could be calculated from the integrals between vibrational ground states. From L_0 and L_1, overlap integrals of the level L_2 could be computed. Each time an overlap integral is calculated, the corresponding probability of transition is confronted to a threshold. If it is lower, the overlap integral is discarded. As a consequence, after having reached a maximum, the number of overlap integrals in higher levels will gradually drop until no more overlap integrals need to be done.

This approach is independent from the energy bounds of the spectrum and so is more general, depending solely on the studied system. However, its implementation has major drawbacks. A previously mentioned one is the arbitrary value of the threshold. This is especially problematic because of one assumption made by Hazra and Nooijen to calculate the overlap integrals. If a precursor, that is to say an overlap integrals belonging to the right-hand side of equations 2.31 or 2.32, is missing because the corresponding probability of transition was too low, the calculation is still performed without it. In other words, a given overlap integral $\langle \boldsymbol{v}' \mid \boldsymbol{v}'' \rangle$ is computed with the overlap integrals available in the memory and those missing, supposed very small, are neglected. This statement breaks the normalization of the overlap integrals and makes it impossible to control correctly the quality of the simulation and the chosen approximation. As a consequence, while the method offers a way to avoid dependence on the spectrum bounds, the absence of a reliable evaluation method of the approximation makes it hardly viable practically. It should be noted that the problem can be partly overcome by recalculating the missing precursors. However, because only two levels are stored at a time, the calculation of these integrals can be time consuming and drop the efficiency of the method.

The analysis of the transition probabilities offers a way for a consistent simulation of the spectrum whose performance rely solely on the nature of the studied system. However, this can be at the expense of the speed when the upper bound of the spectrum is really low with respect to the energy of the transitions between vibrational ground states. Approximation methods done in real time are difficult to handle from the perspective of the storage as there is no way to know beforehand the number of overlap integrals that will need to be computed and saved. In this kind of situation, binary trees are often preferred because of its sturdy organization. While Schumm *et al.* [80] used an algorithm of energy counting with an hash table, they had to take into account the risk of collisions and oversize their storage array, thus losing one of the main advantages of this method over the tree structure.

The overlap integrals that need to be calculated can be known beforehand by applying an evaluation method. Most of them attempt to assess the probabilities of transitions and fix a threshold to choose the most likely ones. However, an efficient *a priori* evaluation is difficult to design because it must meet several conditions. The first one is that it must be relatively fast, that is to say that its computational cost must be kept as affordable as possible. A direct consequence is that the complete calculation of the Franck-Condon integrals is typically out of question. Hence, an

approximation of the transition probabilities is favored. This necessary approximation will cause additional constraints on the evaluation method. The main one is that it must be consistent and efficient whatever the molecular system studied and the approximation of the electronic transition dipole moment applied, with respect to its Taylor series given in equation 1.29. This condition is especially difficult to meet as taking into account all possible cases will often lead to a more complex evaluation system with a higher computation cost. General-purpose evaluation methods are fairly recent and were designed to deal with newly accessible simulations of UV-visible spectra for medium-to-large systems. We will present two methods revolving around the approximation of the Duschinsky rotation [70, 81] and our own procedure [19, 82].

A first and simple approximation is to neglect the mode mixing and consider a one-to-one relation between the modes of the initial and final states. After possible reordering, the Duschinsky transformation matrix **J** corresponds to the identity matrix. The interest of this approximation, called parallel mode approximation, is that the multi-dimensional Franck-Condon integrals can be calculated as products of one-dimensional integrals, using the following relation:

$$\langle \boldsymbol{v}' \mid \boldsymbol{v}'' \rangle = \prod_{i=1}^{N} \langle v_i' \mid v_i'' \rangle \tag{2.34}$$

The one-dimensional Franck-Condon integrals can be calculated using the analytic [34, 35] or recursion formulae [37] for monodimensional oscillators.

Applied to most real systems, a first difficulty arises from the fact that the normal modes are rarely uncoupled and the one-to-one correspondance between the normal modes of the initial and final states does not exist. Ervin *et al.* [81, 87] proposed as a first approximation, to consider the greatest overlap between each mode of the initial state and the final state. This overlap can be qualitatively determined from the Duschinsky matrix, taking the element with the highest absolute value in each row or column.

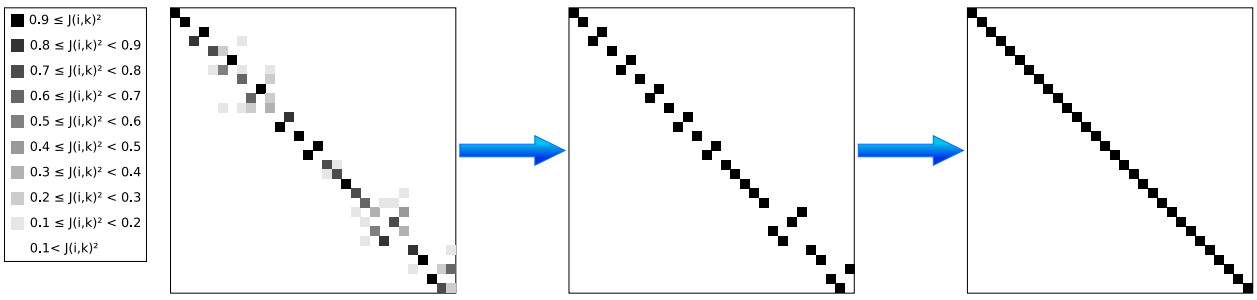


Figure 2.4: Transformation of the correct Duschinsky matrix in an identity matrix in the case of the $\pi \rightarrow \pi^\star$ transition of the phenoxyl radical (DFT - B3LYP/TZVP). In each row, only the element with the highest absolute value is kept and the other one are disregarded. The intensity of the shading corresponds to the relative intensity of the square of each element. For the first two matrices, modes are ordered by increasing frequencies.

Figure 2.4 shows similar procedure for the case of the weakly-allowed $\pi \rightarrow \pi^\star$ transition between electronic states $^2B_1$ and $^2A_2$ of the phenoxyl radical. Calculations have been performed at the Density Functional Theory [88, 89] (DFT) level for the initial state and Time-Dependent Density Functional Theory [13, 90, 91] (TD-DFT) level for the final state, with the B3LYP exchange-

correlation functional and the TZVP basis set. Instead of absolute values, the relative intensities of the square of each element of the rotation matrix **J** are graphically shown in the left drawing. For each column, the highest value is taken and the other ones are disregarded. Practically, it corresponds to setting to one the square of the corresponding element and to zero the other ones. The result is shown in the second matrix. Finally, the modes are reordered and the rotation matrix is equivalent to the identity matrix.

Comparing the correct Duschinsky matrix shown on the left of figure 2.4 and the "new" rotation obtained by this procedure, it is obvious that the uncoupling is generally too strong an approximation. In addition to being often difficult to justify, it will generally lead to important and uncontrollable errors. Moreover, another problem immediately pointed out by Ervin *et al.* [81] is that the parallel mode approximation can simply fail in some cases. It can happen than for a given mode of the initial state, several possible overlaps are found with the selection scheme described previously while for another one, there is none. From the perspective of the rotation matrix, it will mean finding some rows with two or more non-zero elements and others with none.

As an example, let us consider the case of the $S_0(^1A_1) \rightarrow S_1(^1B_2)$ transition[e] of anisole. Calculations were done at the DFT and TD-DFT levels, respectively, using the B3LYP exchange-correlation functional and the 6-311+G(d,p) basis set. In the same way as the phenoxyl radical, we choose the element with the highest absolute value. The result of this approximation is shown in figure 2.5.

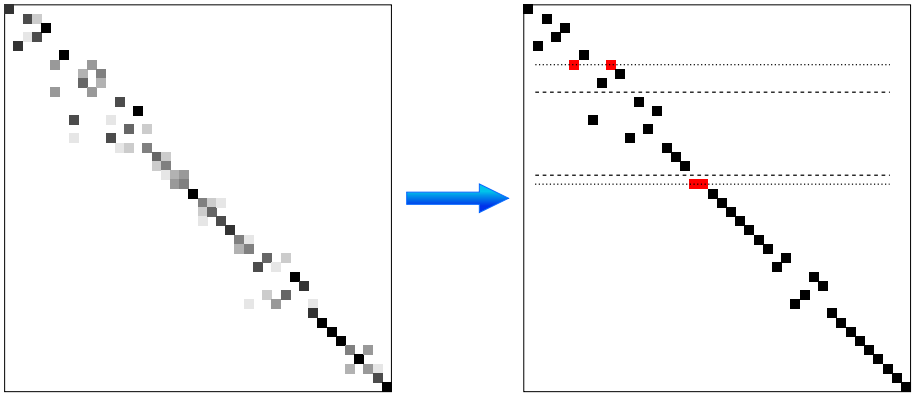


Figure 2.5: Approximation of the Duschinsky matrix by only taking into account the element of highest absolute value in each column in the case of the $S_0 \rightarrow S_1$ (DFT - B3LYP/6-311+G(d,p)). In the right matrix, the dotted lines show the rows with more than one elements and the dashed line the empty rows.

To overcome such a problem, Ervin *et al.* suggested a manual reassignment. This solution is unfortunately not satisfactory for an automated procedure and the arbitrariness of the selection remains. As a result, a treatment of the rotation matrix purely restricted to the parallel mode approximation level is not sustainable for a general-purpose procedure.

Ervin *et al.* [81] proposed a compromising method between a complete treatment of the modes mixing, using the correct rotation matrix, and their neglect as in the parallel mode approximation. They divide the modes into two groups, uncoupled modes and coupled modes. The modes belonging to the first group are treated with the parallel mode approximation while the modes belonging to the second one are treated exactly. In a practical way, the Duschinsky matrix is treated as a block-diagonal matrix instead of an identity matrix.

[e] Transition has a mixed $n \rightarrow \pi^*/\pi \rightarrow \pi^*$ character.

While correcting the flaws we mentioned earlier, this method arises new difficulties. The first one is the criterion to decide if a mode is uncoupled or not. Ervin *et al.* proposed a threshold on the absolute value, considering a mode uncoupled if an element of the corresponding column (or row) represents at least 95% of the sum of all the elements of the column. In addition to the arbitrary criterion, the high value means that if the coupling is very strong between a large number of elements, the gain of this method with respect to the complete treatment of the Duschinsky matrix might be relatively low. However, the method can be really effective with symmetrical rigid systems where the coupling of the modes is limited.

A more subtle problem can occur if one wishes to account for a better approximation of the electronic transition dipole moment beyond the Franck-Condon principle. In this case, terms outside the blocks are necessary to calculate the transition dipole moment integrals and the block-diagonalization can introduce important errors. Indeed, in the case of dispole-forbidden and weakly-allowed transitions, the overlap integral $\langle \boldsymbol{v}' \mid \boldsymbol{v}'' \rangle$ is not the main component in the expression given in equation A.1.

Finally, the difficulty to use this method lies also in the fact that there is no *a priori* evaluation of the impact of these approximations on the outcome results. To be more precise, the method does not include an evaluation of the probability of transitions between each vibrational state. As mentioned in their work, Ervin *et al.* focused on acceleration of the computations rather than on a rigorous prescreening of the overlap integrals. The limitation of the overlap integrals is done through the energy counting and an evaluation of the transition probability in a similar way as Hazra et Nooijen but without cancelling any term of the recursion formulae.

Dierksen and Grimme [70] proposed also a block-diagonalization but choosing a different approach to proceed. Firstly, instead of simply discarding the elements outside the block, they calculate a block-diagonal model rotation matrix by replacing the exact normal modes of the final state by an approximate set. Their procedure requires a threshold on the sum of the elements ${J_{ij}}^2$ for each row and column. This threshold is used to choose the blocks in the original Duschinsky matrix and the new transformation matrix $\tilde{\mathbf{L}}''$ is generated so that out-of-blocks elements are cancelled. Contrary to Ervin *et al.* [81], the new Duschinsky matrix obtained in such a way is not used for the actual calculations but as a prescreening. For each block of this matrix, the Franck-Condon integrals with the corresponding transition energies satisfying the conditions fixed by the bounds of the spectrum are calculated using the recursion formulae of Doktorov *et al.* [65]. The overlap integrals above a second threshold are kept while the other ones are discarded. The "complete" Franck-Condon integrals are obtained by multiplying those preserved in each block and compared to a third threshold. If they are above this lower limit, the correct overlap integral between the same combinations bands is calculated with the original Duschinsky matrix.

With respect to the previously discussed method, this approach improves the reliability of the simulated spectrum by taking into account the correct Duschinsky effect. However, the method suffers from a few flaws, some of them being the same as those pointed out above. The first one is the use of two arbitrary thresholds, the second one generating the two limits presented above to discard the overlap integrals. Because they need to be set by the user, the method is not directly usable in an automatic procedure. Moreover, as pointed out by Jankowiak *et al.* [83], more than

one calculation should be run with different thresholds for the block approximation and the overlap integrals to control the reliability of the prescreening.

The major limitation of this method, in the same way as the one presented by Ervin *et al.* [81], remains its dependency on the distribution of the values ${J_{ij}}^2$ for each row and column of the Duschinsky matrix. The optimum configuration is that only a few elements have a noticeable square value in each case and all the other ones are negligible. After reordering so that these elements are centered around the diagonal, it is possible to easily form small blocks Unfortunately, if the formed blocks are relatively big with respect to the dimension of the Duschinsky matrix, the gain can be seriously reduced. This situation emerges when a large number of elements in each row and column have a small value but none properly stands out. A workaround in this situation would be to lower the first threshold, but the consequence is that the model rotation matrix differs strongly from the original one.

The performances of the approach proposed by Dierksen and Grimme is also strongly intertwined to their choice of storage discussed in the previous section where the binary tree is flushed whenever full. In the case of a restricted available memory and large blocks, recalculations can be important, aggravating the loss in performance. As a consequence, the method seems more suitable to rigid and symmetrical systems [92, 93].

Finally, the same problem of the reliability of evaluating the Franck-Condon integrals in the case of weakly-allowed or dipole-forbidden transitions exists here.

The two methods represent globally an improvement in the computational time needed to simulate a theoretical spectrum. However, efficiency is strongly dependent on the rotation matrix and the coupling of the normal modes. Moreover, the impact of the value of the thresholds on the accuracy of the simulation is difficult to ascertain before the complete calculation has been run.

2.8 The evaluation method *FCClasses*

The procedure we wish to implement should not require much interaction from the user and only a relatively basic knowledge on the theoretical background underlying it to be accessible to non-specialists. The implementation of the method proposed by Dierksen and Grimme is not unfeasible as a "black-box" procedure but a correct automatic choice of the thresholds is difficult without at least an additional run of the calculations with different parameters for the evaluation to control the quality of the first try. This can lead to rather poor performances when the blocks of the model Duschinsky matrix are rather large.

In this work, we looked for a more versatile approach and used the transition probabilities to screen the Franck-Condon integrals. The theoretical basis underlying our method has been originally presented in references [19, 82, 94]. It has been initially implemented in a standalone program, FCCLASSES [95], which was used to describe the principles of the evaluation. We present here its important features.

Firstly, we associate to each overlap integral a class $\mathcal{C}_i$ corresponding to the number of excited modes in the final state. Presented differently, the class $\mathcal{C}_i$ will "contain" all overlap integrals with i non-zero quantum numbers in the final state. This partition is very important in our method since it allows a very straightforward control of the calculations to limit the treatment, and consequently

the computational costs, to the required accuracy (see section 4.4 for a discussion of the influence of the classes on the generated spectrum). Each class contains in theory an infinity of overlap integrals but only a finite number of transitions have a significant probability. Our method estimates for each mode a maximum quantum number $v_{i_{\max}}$ above which the corresponding overlap integrals can be neglected. To do so, reference data from the first two classes, $\mathcal{C}_1$ and $\mathcal{C}_2$, are stored and used to evaluate the overlap to compute in each "higher" class. The advantages of using the overlap integrals of these classes are twofold. The first one is that these integrals are computationally cheap and are generated quickly even in the case of large molecules. The second interest lies in the recursion formulae "adapted" to these classes and the information they provide. As we defined the class with respect to the final state, we will focus on equation 2.31, supposing that the transitions start from the vibrational ground state of the initial state. In the case of $\mathcal{C}_1$, this equation can be reduced to the following expression:

$$\langle\, \mathbf{0}' \mid \boldsymbol{v}'' \,\rangle = \frac{1}{\sqrt{2v_i''}}\Big[\mathrm{D}_i\langle\, \mathbf{0}' \mid \boldsymbol{v}'' - 1_i'' \,\rangle + \sqrt{2(v_i''-1)}\ \mathrm{C}_{ii}\langle\, \mathbf{0}' \mid \boldsymbol{v}'' - 2_i'' \,\rangle\Big] \tag{2.35}$$

In the previous equation, the factors C_{ii} and D_i give respectively information on the effect of the shifts in equilibrium positions and the frequencies on the overlap integrals of overtone, and more precisely on the vibrational progression of mode i. These properties can be verified from the expressions of the matrix $\mathbf{C}$ and vector $\mathbf{D}$ given in equations 2.9 and 2.10, considering the particular case $\mathbf{J} = \mathbf{I}$. Taking into account this approximation, their formulae can be written:

$$\mathbf{C} = 2\boldsymbol{\Gamma}''^{1/2}(\boldsymbol{\Gamma}' + \boldsymbol{\Gamma}'')^{-1}\boldsymbol{\Gamma}''^{1/2} - \mathbf{I} \quad ; \quad \mathbf{D} = -2\boldsymbol{\Gamma}''^{1/2}(\boldsymbol{\Gamma}' + \boldsymbol{\Gamma}'')^{-1}\boldsymbol{\Gamma}''^{1/2}\mathbf{K}$$

The second term of equation 2.35 is then cancelled if $\Gamma'_{ii} = \Gamma''_{ii}$ while the first term is dependent on $\mathbf{K}$.

In the case of overlap integrals belonging to the class $\mathcal{C}_2$, the recursion formula contains also the off-diagonal terms of $\mathbf{C}$.

$$\begin{aligned}\langle\, \mathbf{0}' \mid \boldsymbol{v}'' \,\rangle = \frac{1}{\sqrt{2v_i''}}\Big[&\mathrm{D}_i\langle\, \mathbf{0}' \mid \boldsymbol{v}'' - 1_i'' \,\rangle + \sqrt{2(v_i''-1)}\ \mathrm{C}_{ii}\langle\, \mathbf{0}' \mid \boldsymbol{v}'' - 2_i'' \,\rangle \\ &+ \sqrt{2v_j''}\ \mathrm{C}_{ij}\langle\, \mathbf{0}' \mid \boldsymbol{v}'' - 1_i'' - 1_j'' \,\rangle\Big]\end{aligned} \tag{2.36}$$

where $j \neq i$.

It is straightforward to see that the elements C_{ij} are null in the case where there is no rotation of the normal coordinates ($\mathbf{J} = \mathbf{I}$) and the third term of equation 2.36 will exist only if there is a coupling of modes i and j.

If we store directly the Franck-Condon factors, we will have redundant information on the shifts in equilibrium geometries and frequencies. A simple scheme is used to remove, as far as possible, these effects from the data stored in class $\mathcal{C}_2$.

Before continuing, let us define two arrays $F_{\mathcal{C}_1}$ and $F_{\mathcal{C}_2}$ which store the reference data from classes $\mathcal{C}_1$ and $\mathcal{C}_2$. $F_{\mathcal{C}_1}$ is a bidimensional array containing the Franck-Condon factors such as $F_{\mathcal{C}_1}(i, v_i'') = |\langle\, \mathbf{0}' \mid \mathbf{0}'' + v_i'' \,\rangle|^2$. $F_{\mathcal{C}_2}$ is a tridimensional array whose indexes are both excited modes i and j as well as their quantum number. However, we only store data corresponding to the cases $v_i'' = v_j''$ so only one value is needed. As explained before, the data from class $\mathcal{C}_2$ must essentially

account for the Duschinsky mixing. Hence, instead of the Franck-Condon factors, we keep the quantity $F_{\mathcal{C}_2}(i,j,v_i''=v_j'') = \left|\langle\, \mathbf{0}' \mid \mathbf{0}'' + v_i'' + v_j'' \,\rangle\right|^2 - F_{\mathcal{C}_1}(i,v_i'') \times F_{\mathcal{C}_1}(j,v_i'')/\left|\langle\, \mathbf{0}' \mid \mathbf{0}'' \,\rangle\right|^2$. To explain the choice of this definition, let us assume that there is no mode mixing ($\mathbf{J} = \mathbf{I}$). In this case, the multi-dimensional Franck-Condon integrals can be defined as a product of one-dimensional integral as shown in equation 2.34. As a result, we can write:

$$\langle\, \mathbf{0}' \mid \mathbf{0}'' + v_i'' + v_j'' \,\rangle = \prod_{k=1}^{N} \langle\, 0_k' \mid v_k'' \,\rangle = \langle\, 0_i' \mid v_i'' \,\rangle \times \langle\, 0_j' \mid v_j'' \,\rangle \times \prod_{\substack{k=1 \\ k\neq i,j}}^{N} \langle\, 0_k' \mid 0_k'' \,\rangle$$

A similar expression can be formulated for the overlap integrals belonging to class $\mathcal{C}_1$:

$$\langle\, \mathbf{0}' \mid \mathbf{0}'' + v_i'' \,\rangle = \langle\, 0_i' \mid v_i'' \,\rangle \times \prod_{\substack{k=1 \\ k\neq i}}^{N} \langle\, 0_k' \mid 0_k'' \,\rangle$$

By definition, $F_{\mathcal{C}_2}(i,j,v_i''=v_j'') = 0$ if there is no mode-mixing. To meet this condition, it is enough to write the relation:

$$F_{\mathcal{C}_2}(i,j,v_i''=v_j'') = \left|\langle\, \mathbf{0}' \mid \mathbf{0}'' + v_i'' + v_j'' \,\rangle\right|^2 - \left|\langle\, 0_i' \mid v_i'' \,\rangle\right|^2 \times \left|\langle\, 0_j' \mid v_j'' \,\rangle\right|^2 \times \prod_{\substack{k=1 \\ k\neq i,j}}^{N} \left|\langle\, 0_k' \mid 0_k'' \,\rangle\right|^2$$

The second term in the right-hand side of the previous equation can be written as a product of $\langle\, \mathbf{0}' \mid \mathbf{0}'' + v_i'' \,\rangle$ and $\langle\, \mathbf{0}' \mid \mathbf{0}'' + v_j'' \,\rangle$, noticing that:

$$\begin{aligned}
&\left|\langle\, \mathbf{0}' \mid \mathbf{0}'' + v_i'' \,\rangle\right|^2 \left|\langle\, \mathbf{0}' \mid \mathbf{0}'' + v_j'' \,\rangle\right|^2 \\
&\quad = \left|\langle\, 0_i' \mid v_i'' \,\rangle\right|^2 \times \left\{ \prod_{\substack{k=1 \\ k\neq i}}^{N} \left|\langle\, 0_k' \mid 0_k'' \,\rangle\right|^2 \right\} \times \left|\langle\, 0_j' \mid v_j'' \,\rangle\right|^2 \times \left\{ \prod_{\substack{k=1 \\ k\neq j}}^{N} \left|\langle\, 0_k' \mid 0_k'' \,\rangle\right|^2 \right\} \\
&\quad = \left|\langle\, 0_i' \mid v_i'' \,\rangle\right|^2 \times \left|\langle\, 0_j' \mid v_j'' \,\rangle\right|^2 \times \left\{ \prod_{k=1}^{N} \left|\langle\, 0_k' \mid 0_k'' \,\rangle\right|^2 \right\} \times \prod_{\substack{k=1 \\ k\neq i,j}}^{N} \left|\langle\, 0_k' \mid 0_k'' \,\rangle\right|^2 \\
&\quad = \left|\langle\, 0_i' \mid v_i'' \,\rangle\right|^2 \times \left|\langle\, 0_j' \mid v_j'' \,\rangle\right|^2 \times \left\{ \prod_{\substack{k=1 \\ k\neq i,j}}^{N} \left|\langle\, 0_k' \mid 0_k'' \,\rangle\right|^2 \right\} \times \left|\langle\, \mathbf{0}' \mid \mathbf{0}'' \,\rangle\right|^2
\end{aligned}$$

Hence, we have:

$$\left|\langle\, 0_i' \mid v_i'' \,\rangle\right|^2 \times \left|\langle\, 0_j' \mid v_j'' \,\rangle\right|^2 \times \prod_{\substack{k=1 \\ k\neq i,j}}^{N} \left|\langle\, 0_k' \mid 0_k'' \,\rangle\right|^2 = \frac{\left|\langle\, \mathbf{0}' \mid \mathbf{0}'' + v_i'' \,\rangle\right|^2 \left|\langle\, \mathbf{0}' \mid \mathbf{0}'' + v_j'' \,\rangle\right|^2}{\left|\langle\, \mathbf{0}' \mid \mathbf{0}'' \,\rangle\right|^2}$$

Consequently, we obtain the following relation:

$$F_{\mathcal{C}_2}(i,j,v_i''=v_j'') = \left|\langle\, \mathbf{0}' \mid \mathbf{0}'' + v_i'' + v_j'' \,\rangle\right|^2 - \frac{F_{\mathcal{C}_1}(i,v_i'') \times F_{\mathcal{C}_1}(j,v_i'')}{\left|\langle\, \mathbf{0}' \mid \mathbf{0}'' \,\rangle\right|^2}$$

Given the explanations above, a difficulty still remains to start the prescreening. Because $F_{\mathcal{C}_1}$

and $F_{\mathcal{C}_2}$ will be used to determine the maximum quanta in each class, it means that we need to find a different method to choose these limits for the classes $\mathcal{C}_1$ and $\mathcal{C}_2$. Practically, we set manually the maximum number of quanta in these classes. Since these overlap integrals are computationally cheap, we can set a high enough limit to work in most cases. This somewhat arbitrary choice has been deducted from experience and proved to be valide in most cases. It should be pointed out that more complex schemes could be devised, such as a single threshold on the probability of transitions. The calculations would be run twice for each class, the first time to find the highest quantum number $\mathcal{C}_{1_{\max}} = \max(v_i'')$ and the second time to store the data for the prescreening using this limit for each mode. However, such a scheme, more expensive in terms of computational time, is not necessary in practice and $\mathcal{C}_{1_{\max}}$ and $\mathcal{C}_{2_{\max}}$ are directly set, automatically or by the user.

Calculations are carried out for class $\mathcal{C}_1$ and then class $\mathcal{C}_2$. For class $\mathcal{C}_3$ and above, a two-step procedure is used. Firstly, the maximum quantum numbers $\boldsymbol{v}''_{\max}$ are set by the evaluation method and then the actual calculations are performed. To perform the prescreening, we set an additional limit, $N_I^{\max}$, which represents the maximum number of integrals to compute in a given class. This number lets us control roughly the computational cost of the overall simulation, but at the risk of a loss in accuracy of the generated spectrum if it is set too low. A good compromise between resources usage and precision of the spectrum is set by default. Two thresholds, ϵ_1 and ϵ_2 are defined for the sets of data $F_{\mathcal{C}_1}$ and $F_{\mathcal{C}_2}$, respectively. These two sets are treated in parallel. For each mode i, two maximum quantum numbers, respectively $v_{i_{\max}}^{\mathcal{C}_1}$ and $v_{i_{\max}}^{\mathcal{C}_2}$, are obtained by decrementing the number of quanta v_i'' until the conditions $F_{\mathcal{C}_1}(i, v_i'') \geq \epsilon_1$ for the first set and $F_{\mathcal{C}_2}(i, j, v_i'') \geq \epsilon_2$ with $j \neq i$ for the second set, are met.

The maximum number of quanta is chosen as the maximum of these two values,

$$v''_{i_{\max}} \;=\; \max(v_{i_{\max}}^{\mathcal{C}_1}, v_{i_{\max}}^{\mathcal{C}_2})$$

Once the set $\boldsymbol{v}''_{\max}$ has been defined, the corresponding number of integrals to calculate is roughly estimated, for a given class $\mathcal{C}_n$, as $N_I = {}_NC_n \times \langle \boldsymbol{v}''_{\max} \rangle^n$, where ${}_NC_n$ represents the number of combinations of the n excited oscillators and $\langle \boldsymbol{v}''_{\max} \rangle$ is the arithmetic mean of the N maximum quantum numbers. If the number of integrals to compute N_I is higher than the allowed limit $N_I^{\max}$, the thresholds ϵ_1 and ϵ_2 are increased and the set of maximum quantum numbers $\boldsymbol{v}''_{\max}$ is re-estimated. When the condition $N_I \leq N_I^{\max}$ is fulfilled, all Franck-Condon integrals are computed using the correct maximum number of quanta $v''_{i_{\max}}$ for each mode.

The tests are sufficiently fast to allow a rather large number of trials in a very short computational time. Hence, the thresholds ϵ_1 and ϵ_2 can be chosen very low. By default, they are set to 10^{-9}.

To avoid large needs in storage, the overlap integrals are treated by groups of combination bands of the final state. A group in this case contains all the transitions to final states with the same set of excited modes ($v_i'' > 0$). This strategy matches our storage policy described in section 2.6 and the storage is then kept to a minimum, the storage array being reset after each "group" has been treated. While this causes some recalculations of overlap integrals when progressing to higher modes, the induced computational times on modern computers are generally on a par with those needed for managing large storage requirements.

The method can be straightforwardly used for temperature as described in Appendix C in reference [94]. The temperature was not treated during the development and the implementation of the procedure so it will not be extensively treated further in this document.

2.8.1 Extending the evaluation method to the Herzberg-Teller calculations

The evaluation approach presented above was originally designed to treat calculations done in the frame of the Franck-Condon principle, considering a constant electronic transition dipole moment. The data stored in $F_{\mathcal{C}_1}$ and $F_{\mathcal{C}_2}$ provide general information on the influence of the shift in geometry positions, the change in frequencies and the Duschinsky mixing on the vibrational progression so they are needed whatever the level of approximation used to calculate the intensity of a transition. However, when dealing with weakly-allowed or dipole-forbirdden transitions, we cannot correlate anymore the quantities used for the prescreening with the probabilities of transition. If we develop the electronic transition dipole moment $\boldsymbol{\mu}_{if}$ in a Taylor series up to the first order, the transition dipole moment integral is given by the formula:

$$\begin{aligned}\langle\, \Psi' \mid \boldsymbol{\mu} \mid \Psi'' \,\rangle = \boldsymbol{\mu}_{if}(\mathbf{Q}_0'') \langle\, \boldsymbol{v}' \mid \boldsymbol{v}'' \,\rangle + \sum_{k=1}^{N} \left(\frac{\partial \boldsymbol{\mu}_{if}}{\partial Q_k''}\right)_0 \sqrt{\frac{\hbar}{2\omega_k''}} \Big[\sqrt{v_k''} \langle\, \boldsymbol{v}' \mid \boldsymbol{v}'' - 1_k'' \,\rangle \\ + \sqrt{v_k''+1} \langle\, \boldsymbol{v}' \mid \boldsymbol{v}'' + 1_k'' \,\rangle\Big]\end{aligned} \tag{2.37}$$

The evaluation method we described previously gives all necessary information on the first term of the right-hand side of equation 2.37 but does not take into account the second and third term. In the case of a dipole-forbidden transition, we have $\boldsymbol{\mu}_{if}(\mathbf{Q}_0'') = \mathbf{0}$ and only the second and third terms contribute to the line intensity. Hence, the information provided by $F_{\mathcal{C}_1}$ and $F_{\mathcal{C}_2}$ can fail in accounting for the transition probabilities.

This problem is amplificated if we consider a borderline case where the normal modes are uncoupled and there is no change in equilibrium geometries and frequencies between the initial and final state. In a concise, matrix notation, the above conditions can be summed up as $\mathbf{J} = \mathbf{I}$, $\mathbf{K} = \mathbf{0}$ and $\boldsymbol{\Gamma}' = \boldsymbol{\Gamma}''$. As discussed previously, these conditions will cause the data in $F_{\mathcal{C}_1}$ and $F_{\mathcal{C}_2}$ to be null. As a consequence, the maximum quantum numbers are all equal to zero when doing the evaluation. However, the transitions to final states corresponding to $\mid \mathbf{0}'' + 1_k'' \,\rangle$ have a non-zero intensity because of the second term in the right-hand side of equation 2.37 but would not be taken into account because the prescreening evaluated only zero maximum quantum numbers.

While this hypothetical case is practically unlikely, this example shows the need to improve our evaluation strategy to take into account the influence of the terms stemming from the Herzberg-Teller approximation of the electronic transition dipole moment. To do so, let us define a new array $H_{\mathcal{C}_1}$ which will store the square of the pure Herzberg-Teller contribution for a given mode k for the transition $\langle\, \boldsymbol{v}' \mid \mathbf{0}'' + v_k'' \,\rangle$. These data are calculated contemporarily to those of $F_{\mathcal{C}_1}$ in class $\mathcal{C}_1$:

$$H_{\mathcal{C}_1}(k, v_k'') = \left| \left(\frac{\partial \boldsymbol{\mu}_{if}}{\partial Q_k''}\right)_0 \sqrt{\frac{\hbar}{2\omega_k''}} \left[\sqrt{v_k''} \langle\, \boldsymbol{v}' \mid \mathbf{0}'' + v_k'' - 1_k'' \,\rangle + \sqrt{v_k''+1} \langle\, \boldsymbol{v}' \mid \mathbf{0} + v_k'' + 1_k'' \,\rangle\right] \right|^2$$

Additionally, if we refer to equation 2.37, another difficulty arises from the fact that the summation runs on the N vibrational modes. While this is not a problem for the second term in

the right-hand side of this equation, the third term, where a quantum is created implies that the calculations of $\langle\, \Psi' \mid \boldsymbol{\mu} \mid \Psi'' \,\rangle$ will require to treat the case of modes outside the class we are treating. However, by definition of the class, the modes outside of it can only be excited with a single quantum number in the case of $\langle\, \boldsymbol{v}' \mid \mathbf{0} + v''_k + 1''_k \,\rangle$. Let us consider such a mode k and apply the recursion formula given in equation 2.31 to express the corresponding integral with respect to overlap integrals with lower quanta. As we did in section 2.8, we neglect the effect of the temperature, that is to say we treat all transitions from the vibrational ground state. The recursion formula can then be written:

$$\langle\, \mathbf{0}' \mid \boldsymbol{v}'' + 1''_k \,\rangle = \frac{1}{\sqrt{2(v''_k + 1)}} \Big[\mathrm{D}_k \langle\, \mathbf{0}' \mid \boldsymbol{v}'' \,\rangle + \sum_{l=1}^{N} \sqrt{2v''_l}\, \mathrm{C}_{kl} \langle\, \mathbf{0}' \mid \boldsymbol{v}'' - 1''_l \,\rangle \Big] \tag{2.38}$$

The terms are non-zero only for the modes l belonging to the class.

However, for our evaluation to be as thorough as possible, it needs to account for these contributions outside the class. Consequently, we create a fourth array for our prescreening, $H^{\mathrm{oc}}_{\mathcal{C}_1}$, whose elements are defined, in class $\mathcal{C}_1$, as :

$$H^{\mathrm{oc}}_{\mathcal{C}_1}(k) = \left| \left(\frac{\partial \boldsymbol{\mu}_{if}}{\partial Q''_k} \right)_0 \sqrt{\frac{\hbar}{2\omega''_k}} \langle\, \boldsymbol{v}' \mid \mathbf{0}'' + 1''_k \,\rangle \right|^2$$

It should be noted that, contrary to the three previous arrays, this one is not used to estimate the maximum quantum numbers $\boldsymbol{v}''_{\max}$ but rather to improve the efficiency of the code by reducing the modes to take into account outside the class.

For these two new arrays, we add two new thresholds, ϵ_1^H and ϵ_2^H, respectively. When the first order ot the Taylor series of the electronic transition dipole moment is taken into account, the procedure can be summed up as follows. Class $\mathcal{C}_1$ is treated first. Elements for the arrays $F_{\mathcal{C}_1}$, $H_{\mathcal{C}_1}$ and $H^{\mathrm{oc}}_{\mathcal{C}_1}$ are calculated and stored in their respective arrays. Class $\mathcal{C}_2$ is then treated and the array $F_{\mathcal{C}_2}$ is filled as described in section 2.8. When considering higher classes $\mathcal{C}_n$, the prescreening is proceeding similarly to the case Franck-Condon. To find the maximum quantum numbers, three tests are run in parallel for each normal mode k, confronting $F_{\mathcal{C}_1}(k, v''_k)$ and ϵ_1, $H_{\mathcal{C}_1}(k, v''_k)$ and ϵ_1^H, and $F_{\mathcal{C}_2}(k, l, v''_k)$ and ϵ_2 for each mode $l \neq k$. Once the tests are finished, their maximum number of quanta, respectively $v^{\mathcal{C}_1}_{i_{\max}}$, $v^{H}_{i_{\max}}$ and $v^{\mathcal{C}_2}_{i_{\max}}$ are confronted. The chosen maximum number of quanta is defined as the highest of these three values. The same simple scheme described in the previous section lets us assess the number of integrals N_I that will be computed with the chosen set $\boldsymbol{v}''_{\max}$. If this number exceeds the maximum allowed number of transitions to calculate, ϵ_1, ϵ_1^H and ϵ_2 are increased and the set of maximum quantum numbers $\boldsymbol{v}''_{\max}$ is re-estimated.

If the selection of out-of class modes is enabled ($\epsilon_2^H \geq 0$), then a loop is run once the evaluation is finished on all the normal modes k to find if they should be treated in the Herzberg-Teller contribution. If $H^{\mathrm{oc}}_{\mathcal{C}_1}(k) \geq \epsilon_2^H$, then the mode k is counted in the summation in equation 2.37, otherwise no.

When the prescreening procedure is complete, the calculations of the different transition dipole integrals are done.

2.9 Evaluating the reliability of the prescreening: the sum rules

When using an evaluation method, it is important to have a precise idea of its reliability. In general, the actual calculation is not performed in the prescreening so that the correct intensity of the discarded transitions is not known. Logically, the efficiency of an evaluation must not only be defined by the gain in computational cost but also by its accuracy.

The latter can be evaluated precisely through analytic sum rules. While this method can also bring information on the shape of the spectral distribution [96], we will use it here to control the convergence of the calculation [19].

Let us consider the summation part in the formulae of the absorption and emission spectra given in equations 1.20 and 1.21. We also set aside the energy term, so that all dependance to the incident energy is discarded. The corresponding quantity, I^{tot}, is defined as:

$$I^{\text{tot}} = \sum_i \rho_i \sum_f \left| \langle \Psi_i \mid \boldsymbol{\mu} \mid \Psi_f \rangle \right|^2 = \sum_i \rho_i \sum_f \langle \Psi_i \mid \boldsymbol{\mu} \mid \Psi_f \rangle \langle \Psi_f \mid \boldsymbol{\mu} \mid \Psi_i \rangle \tag{2.39}$$

Using the relation in equation 1.26 and the orthonormality of the vibrational wave functions, the previous relation can be written:

$$I^{\text{tot}} = \sum_i \rho_i \sum_f \langle \boldsymbol{v}' \mid \boldsymbol{\mu} \mid \boldsymbol{v}'' \rangle \langle \boldsymbol{v}'' \mid \boldsymbol{\mu} \mid \boldsymbol{v}' \rangle = \sum_i \rho_i \langle \boldsymbol{v}' \mid \boldsymbol{\mu}_{if} \cdot \boldsymbol{\mu}_{if} \mid \boldsymbol{v}' \rangle \tag{2.40}$$

In the case where all transitions are taken from the vibrational ground state, that is to say we neglect the temperature, I^{tot} is simply the square of the transition dipole moment integral $\langle \mathbf{0}' \mid |\boldsymbol{\mu}_{if}|^2 \mid \mathbf{0}' \rangle$. In the present section, we will consider the general case of a any vibrational state in the initial state.

In the frame of the Franck-Condon approximation, the electronic dipole moment is a constant which can be taken out of the sum.

$$I^{\text{tot}} = \left| \boldsymbol{\mu}_{if}(\mathbf{Q}_0'') \right|^2 \sum_i \rho_i \langle \boldsymbol{v}' \mid \boldsymbol{v}' \rangle = \left| \boldsymbol{\mu}_{if}(\mathbf{Q}_0'') \right|^2 \tag{2.41}$$

Hence, we will have convergence of the spectrum when the condition $\sum_i \rho_i \sum_f \left| \langle \boldsymbol{v}' \mid \boldsymbol{v}'' \rangle \right|^2 = 1$ is met.

In the framework of the Herzberg-Teller approximation of the electronic transition dipole moment, I^{tot} cannot be so simply calculated. Considering the Taylor expansion of the electronic transition dipole moment given in equation 1.29 and the general equation of I^{tot}, it is straightforward to see that we will need the annihilation and creation operators of the second quantization to obtain the sum rules. However, the development of μ_{if} in the Taylor series about the equilibrium geometry of the final state is not immediate because the annihilation and creation operators obtained this way act on the quantum states of the final state and not the initial state. A simpler expression for I^{tot} can be formulated by using a Taylor series of μ_{if} about the equilibrium of the

initial state instead of the final state, limiting the series to the second order:

$$\boldsymbol{\mu}_{if}(\mathbf{Q}') = \boldsymbol{\mu}_{if}(\mathbf{Q}'_0) + \sum_{k=1}^{N} \left(\frac{\partial \boldsymbol{\mu}_{if}}{\partial Q'_k}\right)_0 Q'_k + \frac{1}{2}\sum_{k=1}^{N}\sum_{l=1}^{N} \left(\frac{\partial^2 \boldsymbol{\mu}_{if}}{\partial Q'_k \partial Q''_l}\right)_0 Q'_k Q'_l \tag{2.42}$$

It is noteworthy that $\boldsymbol{\mu}_{if} = \boldsymbol{\mu}_{if}(\mathbf{Q}') = \boldsymbol{\mu}_{if}(\mathbf{Q}'')$, however the equivalence of each term in the right-hand side of equations 1.29 and 2.42 is not true.

Because the expression of I^{tot} is now dependent on the chosen electronic state as the reference of the Taylor expansion, we will use a subscript to represent this dependence, $I^{\text{tot}}_{\mathbf{Q}'}$ for the initial state and $I^{\text{tot}}_{\mathbf{Q}''}$ for the final state. Using equation 2.42 and limiting the Taylor series to the zeroth and first orders, $I^{\text{tot}}_{\mathbf{Q}'}$ can be written:

$$I^{\text{tot}}_{\mathbf{Q}'} = \sum_i \rho_i \langle\, \boldsymbol{v}' \mid \left[\boldsymbol{\mu}_{if}(\mathbf{Q}'_0) + \sum_{k=1}^{N} \left(\frac{\partial \boldsymbol{\mu}_{if}}{\partial Q'_k}\right)_0 Q'_k\right] \cdot \left[\boldsymbol{\mu}_{if}(\mathbf{Q}'_0) + \sum_{k=1}^{N} \left(\frac{\partial \boldsymbol{\mu}_{if}}{\partial Q'_k}\right)_0 Q'_k\right] \mid \boldsymbol{v}' \,\rangle \tag{2.43}$$

The normal coordinates Q'_i are transformed as a combination of annihilation and creation operators following the relation given in equation 1.53. We recall the properties of these operators, taking into account the orthogonality of the vibrational states:

$$\langle\, \boldsymbol{v}' \mid a_k^\dagger \mid \boldsymbol{v}' \,\rangle = \sqrt{v'_k + 1}\, \langle\, \boldsymbol{v}' \mid \boldsymbol{v}' + 1'_k \,\rangle = 0 \tag{2.44}$$

$$\langle\, \boldsymbol{v}' \mid a_k \mid \boldsymbol{v}' \,\rangle = \sqrt{v'_k}\, \langle\, \boldsymbol{v}' \mid \boldsymbol{v}' - 1'_k \,\rangle = 0 \tag{2.45}$$

Since the vibrational wave functions are orthonormal to each other, the only way to have non-zero terms is by applying the same number of times the annihilation and the creation operators on the same quanta. As a consequence, in the framework of the Herzberg-Teller approximation, we will have non-zero terms in two cases:

$$\langle\, \boldsymbol{v}' \mid a_k^\dagger a_k \mid \boldsymbol{v}' \,\rangle = (v'_k + 1)\, \langle\, \boldsymbol{v}' \mid \boldsymbol{v}' \,\rangle = v'_k + 1 \tag{2.46}$$

$$\langle\, \boldsymbol{v}' \mid a_k a_k^\dagger \mid \boldsymbol{v}' \,\rangle = v'_k\, \langle\, \boldsymbol{v}' \mid \boldsymbol{v}' \,\rangle = v'_k \tag{2.47}$$

From these remarks, it is straightforward to see that all terms with an odd power of any normal coordinate Q'_k will be null. Hence, the terms $\boldsymbol{\mu}_{if} \cdot (\partial \boldsymbol{\mu}_{if}/\partial Q'_k) Q'_k$, as well as the terms $(\partial \boldsymbol{\mu}_{if}/\partial Q'_k) Q'_k \cdot (\partial \boldsymbol{\mu}_{if}/\partial Q'_l) Q'_l$ with $l \neq k$ will be null.

To formulate completely the analytic limit, we will express it with respect to the Cartesian components of the electronic transition dipole moment and its derivatives. The Cartesian coordinates x, y, z will be represented by the greek letter τ. Taking into account the previous observations, equation 2.43 can be written

$$I^{\text{tot}}_{\mathbf{Q}'} = \sum_{\tau=x,y,z} \sum_i \rho_i \langle\, \boldsymbol{v}' \mid \mu_{if}(\mathbf{Q}'_0, \tau)^2 + \sum_{k=1}^{N} \left(\frac{\partial \mu_{if}(\tau)}{\partial Q'_k}\right)_0^2 {Q'_k}^2 \mid \boldsymbol{v}' \,\rangle \tag{2.48}$$

with

$${Q'_k}^2 = \frac{\hbar}{2\omega'_k} \times [a_k a_k + a_k a_k^\dagger + a_k^\dagger a_k + a_k^\dagger a_k^\dagger]$$

In the case of the Herzberg-Teller approximation, the analytic limit $I^{\text{tot}}_{\mathbf{Q}'}$ is given by the expression:

$$I^{\text{tot}}_{\mathbf{Q}'} = \sum_{\tau=x,y,z} \mu_{if}(\mathbf{Q}'_0,\tau)^2 + \sum_i \rho_i \sum_{k=1}^{N} \left(\frac{\partial \mu_{if}(\tau)}{\partial Q'_k}\right)_0^2 \frac{\hbar}{2\omega'_k}(2\omega'_k+1) \tag{2.49}$$

A practical problem arises from the fact that we might not have the necessary data for the Taylor series of the electronic transition dipole moment about the equilibrium geometry of the initial state. In this case, it is also necessary to have an analytic limit depending on the final state, $I^{\text{tot}}_{\mathbf{Q}''}$. As previously, we replace the electronic transition dipole moment by its Taylor series about the equilibrium geometry, this time of the final state, given in equation 1.29:

$$I^{\text{tot}}_{\mathbf{Q}''} = \sum_i \rho_i \langle\, \boldsymbol{v}' \mid \left[\boldsymbol{\mu}_{if}(\mathbf{Q}''_0) + \sum_{k=1}^{N}\left(\frac{\partial \boldsymbol{\mu}_{if}}{\partial Q''_k}\right)_0 Q''_k\right] \cdot \left[\boldsymbol{\mu}_{if}(\mathbf{Q}''_0) + \sum_{k=1}^{N}\left(\frac{\partial \boldsymbol{\mu}_{if}}{\partial Q''_k}\right)_0 Q''_k\right] \mid \boldsymbol{v}' \,\rangle \tag{2.50}$$

Let us use the inverse of the Duschinsky transformation given in equation 1.64 to express the normal coordinates of the final state as a linear combination of the normal coordinates of the initial state:

$$\mathbf{Q}'' = \mathbf{J}^{-1}\mathbf{Q}' - \mathbf{J}^{-1}\mathbf{K} \tag{2.51}$$

Introducing equation 2.51 into equation 2.50, we obtain:

$$\begin{aligned} I^{\text{tot}}_{\mathbf{Q}''} &= \sum_i \rho_i \langle\, \boldsymbol{v}' \mid \left[\boldsymbol{\mu}_{if}(\mathbf{Q}''_0) + \sum_{k=1}^{N}\left(\frac{\partial \boldsymbol{\mu}_{if}}{\partial Q''_k}\right)_0 \sum_{l=1}^{N}\left(\mathrm{J}^{\mathrm{T}}\right)_{kl} Q'_l + \left(\mathrm{J}^{\mathrm{T}}\right)_{kl}\mathrm{K}_l\right]^2 \mid \boldsymbol{v}' \,\rangle \\ &= \sum_{\tau=x,y,z}\sum_i \rho_i \langle\, \boldsymbol{v}' \mid \left[\mu_{if}(\mathbf{Q}''_0,\tau) + \sum_{k=1}^{N}\left(\frac{\partial \mu_{if}(\tau)}{\partial Q''_k}\right)_0 \sum_{l=1}^{N}\left(\mathrm{J}^{\mathrm{T}}\right)_{kl} Q'_l + \left(\mathrm{J}^{\mathrm{T}}\right)_{kl}\mathrm{K}_l\right]^2 \mid \boldsymbol{v}' \,\rangle \end{aligned} \tag{2.52}$$

The square of the double sum in the bracket,

$$\sum_{k=1}^{N}\left(\frac{\partial \mu_{if}(\tau)}{\partial Q''_k}\right)_0 \sum_{l=1}^{N}\left(\mathrm{J}^{\mathrm{T}}\right)_{kl} Q'_l + \left(\mathrm{J}^{\mathrm{T}}\right)_{kl}\mathrm{K}_l \,,$$

is then expanded and only the terms with even powers of the normal coordinates Q'_k, including the power 0 are kept. Replacing the remaining normal coordinates by the annihilation and creation operators and following the method presented before, the analytic limit $I^{\text{tot}}_{\mathbf{Q}''}$ can be expressed as:

$$\begin{aligned} I^{\text{tot}}_{\mathbf{Q}''} = \sum_{\tau=x,y,z} \mu_{if}(\mathbf{Q}''_0,\tau)^2 + \sum_i \rho_i \sum_{k=1}^{N}\left(\frac{\partial \mu_{if}(\tau)}{\partial Q''_k}\right)_0^2 \sum_{l=1}^{N}\Bigg\{ \left(\mathrm{J}^{\mathrm{T}}\right)^2_{kl} \frac{\hbar}{2\omega'_l}(2\omega'_l+1) \\ + 2\mu_{if}(\mathbf{Q}''_0,\tau)\left(\mathrm{J}^{\mathrm{T}}\right)_{kl}\mathrm{K}_l + \sum_{m=1}^{N}\left(\mathrm{J}^{\mathrm{T}}\right)_{kl}\mathrm{K}_l\left(\mathrm{J}^{\mathrm{T}}\right)_{km}\mathrm{K}_m \Bigg\} \end{aligned} \tag{2.53}$$

It is important to note that the exactness of the analytic limit $I^{\text{tot}}_{\mathbf{Q}''}$ given above depends on the accuracy of the linear transformation approximation given in equation 2.51. If the Duschinsky transformation does not correctly represent the normal coordinates of the final state with respect to those of the initial state, then the analytic limit can be partially erroneous.

Finally, let us consider the Taylor series of the electronic transition dipole moment about the equilibrium of the geometry of the initial state up to the second order. Because of the complexity of the analytic limit depending on the final state, we will only treat here the case of $I^{\text{tot}}_{\mathbf{Q}'}$. Replacing $\boldsymbol{\mu}_{if}$ in equation 2.40 by its Taylor expansion, one obtains:

$$I^{\text{tot}}_{\mathbf{Q}'} = \sum_i \rho_i \langle\, \boldsymbol{v}' \mid \left[\boldsymbol{\mu}_{if}(\mathbf{Q}'_0) + \sum_{k=1}^{N} \left(\frac{\partial \boldsymbol{\mu}_{if}}{\partial Q'_k} \right)_0 Q'_k + \sum_{k=1}^{N} \sum_{l=1}^{N} \left(\frac{\partial^2 \boldsymbol{\mu}_{if}}{\partial Q'_k \partial Q'_l} \right)_0 Q'_k Q'_l \right]^2 \mid \boldsymbol{v}' \,\rangle \tag{2.54}$$

Firstly, let us consider only the operator in equation 2.54. After expanding it, we only keep terms with even powers of each normal coordinate Q'_k. By doing this, we have the following terms, for each Cartesian coordinate represented by τ:

$$\begin{aligned} &\mu_{if}(\mathbf{Q}'_0,\tau)^2 + \sum_{k=1}^{N} \Bigg\{ \left(\frac{\partial \mu_{if}(\tau)}{\partial Q'_k} \right)_0^2 {Q'_k}^2 + 2\mu_{if}(\tau)(\mathbf{Q}'_0) \left(\frac{\partial^2 \mu_{if}(\tau)}{\partial {Q'_k}^2} \right)_0 {Q'_k}^2 + \left(\frac{\partial^2 \mu_{if}(\tau)}{\partial {Q'_k}^2} \right)_0^2 {Q'_k}^4 \\ &\quad + \sum_{\substack{l=1 \\ l\neq k}}^{N} \left(\frac{\partial^2 \mu_{if}(\tau)}{\partial {Q'_k}^2} \right)_0 \left(\frac{\partial^2 \mu_{if}(\tau)}{\partial {Q'_l}^2} \right)_0 {Q'_k}^2 {Q'_l}^2 + 2 \left(\frac{\partial^2 \mu_{if}(\tau)}{\partial Q'_k \partial Q'_l} \right)_0^2 {Q'_k}^2 {Q'_l}^2 \Bigg\} \end{aligned}$$

Considering the terms above, we have 3 possible combinations of normal coordinates, ${Q'_k}^2$, ${Q'_k}^4$ and ${Q'_k}^2 {Q'_l}^2$. After replacing the normal coordinates by the annihilation and creation operators, it is unnecessary to fully develop the obtained polynomials. As a matter of fact, keeping in mind the orthogonality of the vibrational wave functions, only the operators keeping unchanged the vibrational states, that is to say without creating or annihilating any quantum, are to be taken into account. As a consequence, we finally obtain the following expressions:

$$\langle\, \boldsymbol{v}' \mid {Q'_k}^2 \mid \boldsymbol{v}' \,\rangle = \frac{\hbar}{2\omega'_k}\big(2v'_k + 1\big) \tag{2.55}$$

$$\begin{aligned} \langle\, \boldsymbol{v}' \mid {Q'_k}^2 {Q'_l}^2 \mid \boldsymbol{v}' \,\rangle &= \frac{\hbar^2}{4\omega'_k \omega'_l}\big((2v'_k + 1)(2v'_l + 1)\big) \\ &= \frac{\hbar^2}{4\omega'_k \omega'_l}\big(4v'_k v'_l + 2v'_k + 2v'_l + 1\big) \end{aligned} \tag{2.56}$$

In the case of ${Q'_k}^4$, more combinations of operators are possible. To resolve this problem, let us develop the polynomial $[a_k + a_k^\dagger]^4$:

$$[a_k + a_k^\dagger]^4 = a_k^4 + 4a_k^3 a_k^\dagger + 6a_k^2 {a_k^\dagger}^2 + 4a_k {a_k^\dagger}^3 + {a_k^\dagger}^4$$

It is straightforward to see that only the middle term in the right-hand side of the previous equation, $6a_k^2 {a_k^\dagger}^2$ can keep unchanged the number of quanta in the vibrational state. If we apply this operator to $\boldsymbol{v}'$, we obtain:

$$\begin{aligned} \langle\, \boldsymbol{v}' \mid 6a_k^2 {a_k^\dagger}^2 \mid \boldsymbol{v}' \,\rangle &= \langle\, \boldsymbol{v}' \mid a_k a_k a_k^\dagger a_k^\dagger + a_k a_k^\dagger a_k a_k^\dagger + a_k a_k^\dagger a_k^\dagger a_k + a_k^\dagger a_k a_k a_k^\dagger + a_k^\dagger a_k a_k^\dagger a_k + a_k^\dagger a_k a_k a_k^\dagger \mid \boldsymbol{v}' \,\rangle \\ &= v'_k(v'_k - 1) + {v'_k}^2 + v'_k(v'_k + 1) + (v'_k + 1)(v'_k + 2) + (v'_k + 1)^2 + (v'_k + 1)v'_k \\ &= 6{v'_k}^2 + 6v'_k + 3 \end{aligned}$$

Using the previous relation, the term $\langle\, \boldsymbol{v}' \mid {Q'_k}^4 \mid \boldsymbol{v}' \,\rangle$ can be written:

$$\langle\, \boldsymbol{v}' \mid {Q'_k}^4 \mid \boldsymbol{v}' \,\rangle = \frac{\hbar^2}{4{\omega'_k}^2}\left(6{v'_k}^2 + 6v'_k + 3\right) \tag{2.57}$$

Using equations 2.55, 2.56 and 2.57, the analytic limit $I^{\text{tot}}_{\mathbf{Q}'}$ is given by the following expression:

$$\begin{aligned} I^{\text{tot}}_{\mathbf{Q}'} = \sum_{\tau=x,y,z} \mu_{if}(\mathbf{Q}''_0,\tau)^2 + \sum_i \rho_i \sum_{k=1}^{N} \Bigg\{ & \frac{\hbar}{2\omega'_k}\Bigg[\left(\frac{\partial \mu_{if}(\tau)}{\partial Q'_k}\right)_0^2 \\ & + 2\mu_{if}(\tau)(\mathbf{Q}'_0)\left(\frac{\partial^2 \mu_{if}(\tau)}{\partial {Q'_k}^2}\right)_0\Bigg]\left(2v'_k + 1\right) \\ & + \frac{\hbar^2}{4{\omega'_k}^2}\left(\frac{\partial^2 \mu_{if}(\tau)}{\partial {Q'_k}^2}\right)_0^2 \left(6{v'_k}^2 + 6v'_k + 3\right) \\ & + \sum_{\substack{l=1 \\ l\neq k}}^{N} \frac{\hbar^2}{4\omega'_k\omega'_l}\Bigg[\left(\frac{\partial^2 \mu_{if}(\tau)}{\partial {Q'_k}^2}\right)_0 \left(\frac{\partial^2 \mu_{if}(\tau)}{\partial {Q'_l}^2}\right)_0 \\ & + 2 \times \left(\frac{\partial^2 \mu_{if}(\tau)}{\partial Q'_k \partial Q'_l}\right)_0^2\Bigg]\left(4v'_k v'_l + 2v'_k + 2v'_l + 1\right)\Bigg\} \end{aligned} \tag{2.58}$$

Chapter 3

Implementation inside GAUSSIAN

3.1 Introduction

The previous chapters presented the necessary theoretical background and calculations strategies to compute the Franck-Condon integrals and finally simulate absorption or emission UV-visible spectra. In this chapter, we will present a practical procedure to generate these spectra.

Most available programs to compute Franck-Condon integrals are standalone. Many of them are direct implementations of calculation [45, 68] or prescreening [70, 81, 82] methods. Because they are mainly intended for internal use, they are generally lacking a simple and easy-to-use interface or are not distributed at all. This confers an exclusive character to them and often makes difficult their usage as well as the interpretation of the data they produce. Moreover, being a direct implementation of theoretical methods, the code is in general straightforwardly written, without paying much attention to its optimization. However, considering the steep increase of overlap integrals to compute with respect to the spectral range, (see figure 2.3 for an example of the relation between the spectra range and the number of transitions), the program will need to repeat a huge number of times very similar operations. Consequently, a lack of code efficiency can have a significant impact on the computational costs to generate each spectrum, irrespective of the theoretical performances of the method used to compute it.

In our work, we aimed at providing an efficient and simple procedure to compute spectra of radiative transitions, accessible to non-specialists. Instead of implementing a standalone software, we integrated it into the well-known quantum mechanical computational package GAUSSIAN [18]. We will briefly present this program in the next section, including a description of its internal structure. For now, we will focus on the main objectives and the constraints we fixed or had to follow to code our procedure. We will refer to it inside GAUSSIAN as a module, that we will call FRANCK. As mentioned before, the prescreening method we use her has been originally encoded in a standalone software, FCCLASSES [95]. While the latter contained some remarkable algorithms, a complete rewriting of the implementation with further optimization has been performed here to allow a smooth and efficient integration into GAUSSIAN.

Two main constraints imposed by the integration are the necessity to fully blend into the existing code and carefully manage the memory usage of our module. The first point means adapting to the general framework and the available computational tools to avoid redundant code

and efficiently use the resources and data available through the calculations done by other modules. The memory requirement must also be carefully controlled as our data will have to "coexist" with other information needed by GAUSSIAN inside the allocated memory. This problem is far more complex that for a standalone program as we do not have any direct control on the memory resources.

In addition to these imposed limitations, we also set several supplementary requirements for our code to meet. We chose to write FRANCK in a highly modular form. Avoiding a monolithic structure, it is possible to simply and easily add new computational methods as well as maintaining the code more efficiently. Also, *ad hoc* functions specialized in a particular task can be called only if required by the calculation settings, avoiding real-time controls during calculations that might hamper the efficiency of the procedure. This decision, however, causes a slight redundance of the code.

We also wished for our procedure to be as automated as possible, requiring as little setting as possible. Two objectives were pursued in this case. This choice makes the whole procedure easier to handle even for users with little knowledge of the internal structure and theoretical background of our code and the Franck-Condon calculations in general. While, for an experienced user, this means being able to focus on the settings he or she wished to adjust without having to pay much attention to the other parameters. To reduce to a minimum the necessary interventions from the user, it means that all necessary options to perform the calculations can have a default value adapted to most cases, often stemming from extensive benchmark studies, or that their value can be correctly set by FRANCK with respect to the studied systems. A major problem, discussed in sections 2.7 and 2.8, concerns the settings required by the evaluation methods. The current prescreening method we implemented, described in section 2.8, needs a few parameters but they can be safely set to generic values with satisfactory results in almost every case. Some values, such as the bounds of the simulated spectrum, can be simply adjusted by the program once the energy of the transition between the vibrational ground states of the initial and final states is known. In the following description, we will present the parameters that can be adjusted and their default value.

This choice of simplicity in usage was also made for the output, limited to a "compact", readable form by default. After having summed up the main features chosen for the generation of the spectrum, the procedure lists the intensity of the main transitions as well as the vibrational states involved, and finally prints the spectrum. For the output just as the calculation, many parameters are available to adapt the generation of the spectrum to a particular case which is not satisfactorily treated with the default settings or to fine-tune it for a specific need of the user.

While satisfying these constraints, a particular attention was paid that the chosen solutions would not impede at all - or to a minimum if unavoidable - the efficiency of our calculations. This imposed sometimes to stray from a straightforward encoding of the formulae we described before in chapter 2 and choose more suitable expressions, at the possible expense of their "direct" readibility when controlling the code.

Following the modular structure of the code, this chapter will be organized as follows. The first part will be centered on a general presentation of the integration inside the GAUSSIAN development version (revision version G03). This will be done from a technical point of view with an emphasis

on the way the user can interact with it and control the calculation. Then, the diverse functions of the procedure will be globally presented by tasks, such as the extraction of data, the output or the calculation of the transition dipole moment integrals.

3.2 Overview of GAUSSIAN and the interaction between FRANCK and the rest of the software

Based on the basic laws of quantum mechanics, GAUSSIAN is able to predict the energies, the molecular structures and the vibrational frequencies of molecular systems, as well as many other molecular properties. It is commonly used to study molecules or reactions in a wide range of conditions, even for systems which might be difficult or impossible to study experimentally, such as transition states or compounds with a very short lifetime. The relative simplicity of usage along with its performances and its versatility makes it a very interesting tool for quantum chemists and non-specialists.

From a programming perspective, GAUSSIAN is written in Fortran and its structure is highly modular. While the knowledge of this language is unnecessary to understand the following presentation and the rest of the chapter, it is useful to know that a subroutine is a callable sequence of instructions defined inside a Fortran program[a]. The software relies on a set of independent, specialized subprograms, called links, that can be called if necessary. Three links, Link 0, Link 1 and Link 9999, have a central place in the execution of GAUSSIAN and are always called, whatever the type of calculations required. Link 0 is the GAUSSIAN executable which is launched by the user. It remains as a background task while GAUSSIAN is running. Soon after the initialization, Link 1 is called. This subprogram has an interesting role for us because, among its many tasks, it is in charge of the treatment of the instructions given in input by the user. The reading of those commands and parameters followed by their interpretation leads to the creation of a *route* consisting of an ordered list of links, that will be followed by the program. Finally, the last link, Link 9999 is in charge of closing the calculation. It prints a summary of the results of the calculation in the output and manages the final writing in the different working files of GAUSSIAN.

The complete list of the available links as well as their role would be long and out of the scope of this presentation. However, it should be pointed out that the naming of a link follows an established convention that gives a general idea of its function. Each subprogram, except links 0 and 1 described above, is an executable file called "lxyy.exe" or "l$xxyy$.exe". The number x (or xx) represents the *overlay*, that we could define as the identification number of a group of links which have a similar purpose or aim at determining similar chemical properties. As an example, the *overlay* 5 ($x = 5$) concerns all the links that compute the Self-Consistent Field (SCF) wavefunction, and links of the *overlay* 7 aim at calculating the first and second derivatives of the energy with respect to nuclear coordinates. The number yy identifies uniquely the link inside the overlay.

In this work, we implemented our procedure inside the link l716. This link completes the evaluation of the energy derivatives by computing the harmonic vibrational frequencies and normal modes from the force constants. An interesting feature implemented in this link is the possibility

[a] A subroutine can be seen as a small subprogram intended for a very specific task that needs to be repeated several times.

to read directly from the input file its own option in a specific, dedicated zone. Our module can take advantage of this functionality to offer a large panel of parameters to set. To understand how it is working, let us consider a simple generic input file.

```
%NProcs=Number_of_processors_to_use
%Mem=Memory_allocated_to_GAUSSIAN
%Chk=name_of_the_checkpoint_file
#P route section

Description

Charge Multiplicity
Atomic Types and Coordinates
```

The first two lines set the resources at the disposal of GAUSSIAN, that is to say the number of processors and the memory it is allowed to use. The memory set here is then used as an upper limit that cannot be exceeded by the whole of the variables contemporarily stored in the GAUSSIAN working array. Depending on the type of calculation requested to GAUSSIAN, the program can have very disparate needs in memory. For the same amount of available memory, the procedure FRANCK, when called, can have different resources at its disposal. To be able to work in most conditions and also not hamper the other calculations that remain to be done, our procedure must use as little memory as possible.

The third line instructs GAUSSIAN to store the values it has computed into a binary file called the *checkpoint file*. This file contains quantities stored at machine precision that have been calculated by GAUSSIAN as well as the options given in the input. The data it stores will be of importance later for our computations.

These lines are options for the Link 0 and their order is not important. The rest of the input is interpreted by Link 1 and is partitionned in sections defined as a block of text (instruction, comments, data...) terminated by a blank line.

The line starting by "`#`" is called the route section. This is the part where the user can define the type of calculations GAUSSIAN must perform along with the parameters it has to use. Three important data are generally given there, the type of calculation such as geometry optimization (keyword `Opt`), frequency calculations (keyword `Freq`), the method (Hartree-Fock, DFT, TD-DFT, semi-empirical...), and the basis set if needed by the method. While default values exist for these three fields, they are generally not adapted to the needs of the user and must be explicitly given.

The description, also called in GAUSSIAN the title section can be used by the user to write comments on the calculation. It is not interpreted by the program.

Finally, the specifications of the molecular system are given, starting by the charge and multiplicity on a single line and a list of the atoms and their coordinates afterwards[b].

[b]The molecule specification can also be given as a Z-matrix instead of Cartesian coordinates as described in our example.

To enable the reading of additional options in the input file by the link l716, we add the keywords `VIBROT` and `READANHARM` as suboptions of `Freq`. The link will now look for its own options in a dedicated section at the end of the input file[c] :

```
%NProcs=Number_of_processors_to_use
%Mem=Memory_allocated_to_GAUSSIAN
%Chk=name_of_the_checkpoint_file
#P FREQ=(VIBROT,READANHARM) rest of the route section

Description

Charge Multiplicity
Atomic Types and Coordinates

"dinautil" section
```

The *"dinautil" section* is primarily intended to direct the calculations of spectroscopic properties beyond the harmonic analysis of the rovibrational transitions.

The main options concerning the type of job run by GAUSSIAN and the necessary parameters to carry it out as desired by the user remain in the *route section* at the beginning of the input while the specific settings to simulate the Franck-Condon spectra are in the zone below, that we will refer as the *dinautil section*. It is important to note that keywords cannot be freely interchanged between these two sections. While the first one is interpreted by Link 1, the second one is analyzed by Link 716 and they do not share the same list of recognized keywords. Another remark, consequence of the previous observation, is that the *dinautil section* can only be controlled by l716 and more exactly a specific subroutine inside it, once it has been called. This means that in case of errors in this line, the program can only indicate it after possibly long calculations. Consequently, it is advisable to carefully set the options in the *dinautil section* to avoid such a problem.

Finally, for experienced users of GAUSSIAN that are using the *dinautil section* for the first time, the keywords wrote there are not summed up by GAUSSIAN at the beginning of the output as it is the case for the *route section* because it has not been read by Link 1 when the corresponding part of the output is written. However, hese keywords can be found further in the output:

```
-----------------------------------------------------------------------
 Using the following non-standard input for DiNa:
  "dinautil" section
```

However, this section is in the flow of the output and is not easily pinpointed, so it can be rather cumbersome to find the settings used for the generation of the spectrum. Consequently, the main parameters chosen by the user or set by default are summed up at the beginning of the output written by FRANCK.

[c]We presented the *dinautil section* in the input after the listing of the atomic coordinates but it can be later depending on the options, for example when isotopic substitution, external basis set, specific thermodynamic conditions, etc. are requested by the user. In this case, it is after the listing of these data

The generation of a spectrum requires in practice a single keyword in the *dinautil section* which corresponds to the orders of the Taylor series of the electronic transition dipole moment shown in equation 1.29 that must be taken into account:

- `FC` : computes the Franck-Condon spectrum (order 0 of the Taylor series)
- `FCHT` : computes the Franck-Condon Herzberg-Teller spectrum (order 0 and 1 of the Taylor series)
- `HT` : computes the Herzberg-Teller spectrum (order 1 of the Taylor series)

It should be noted that we use the expression "Herzberg-Teller approximation" to refer generically to the first order of the Taylor expansion of the electronic transition dipole moment independently of the state of reference. A stricter definition should restrict its usage to the case of the initial state chosen as reference to match the original work of Herzberg and Teller [29]. Currently, the procedure uses the equilibrium geometry of the final state for the Taylor expansion. As a result, the HT approximation in this case will be different from the original HT one and corresponds more precisely to a FCHT calculation. A more complete discussion will be done in section 3.9 and in the study of the porphyrin in section 4.7.

From now on, we will focus on the options for the generation of theoretical spectra. The corresponding keywords must all be given in the *dinautil section*, so their position in the input file will not be given in the following discussions. For convenience, we will also adopt the acronyms FC, FCHT and HT to describe the *type of calculation* of the transition dipole moment integrals with respect to the approximation of the electronic transition dipole moment.

While one of the previously given keyword is enough to start the generation of a UV-vis spectrum, a supplementary option can be given to indicate the type of transition to consider.

- `ABS/ABSORPTION` (*default*) : simulates an absorption spectrum
- `EMI/EMISSION` : simulates an emission spectrum

By default, the spectrum of absorption transitions is generated. An emission spectrum can be chosen with the keywords `EMI` or `EMISSION`.

Once the *type of calculation*, and optionally the *type of transition* to compute, has been found, the procedure FRANCK is called. While it is fully integrated into GAUSSIAN and uses, as far as possible, the computational tools and data it offers, our procedure works mainly as a plugin, connected to the subprogram l716 by two subroutines. The first one interprets the keywords given in input for the spectrum computation and controls the vailidity of the values given by the user. It also sets the default values for the missing options and prints the main information in the output. For example, in the case of an FC calculation and an absorption spectrum, the output is:

```
**********************************************************************

            Generation of the Franck-Condon spectrum

**********************************************************************

Approx. of the electronic transition dipole moment: FC
Type of transition requested: ABSORPTION
```

Another important task of this first subroutine is to extract the data needed to perform the calculations. Finally, it calls the second subroutine which will manage the actual computations to generate the final spectrum. These two subroutines constitute the central structure of FRANCK and most tasks are directed by one of them. The diagram in figure 3.1 shows the general structure of our procedure and the task of the main subroutines.

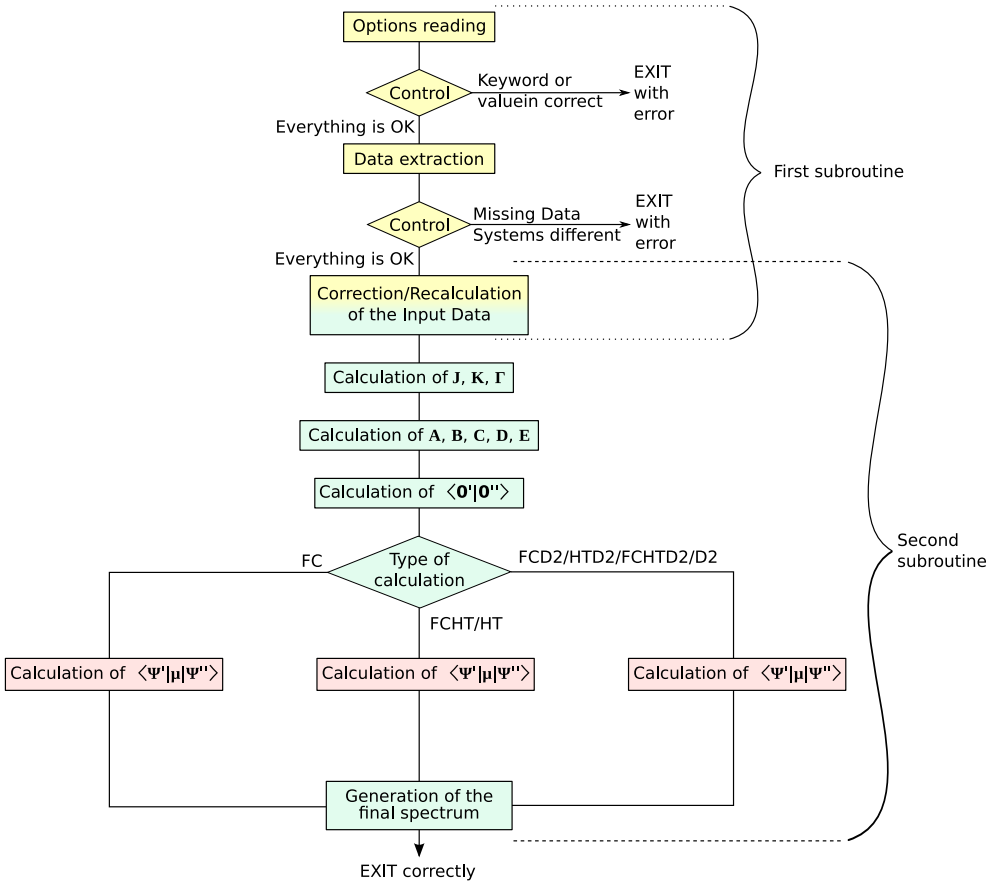


Figure 3.1: Simplified diagram of the organization of FRANCK

The blocks with a bright yellow background correspond to tasks performed by the first subroutine (FRANCK1), while those with a bright blue background are the tasks of the second subroutine (FRANCK2). The correction and recalculation of the input data, described in section 3.3, is partly done by each subroutine, hence the gradient has a background color of the block. The computations of the transition dipole moment integrals are carried out by a specific subroutine depending on the *type of calculation* requested, so a magenta background was used, that we can generically

call FRANCK3.

In this diagram, we also considered the case of calculations taking into account the second order in the Taylor series of the electronic transition dipole moment, referred as D2 in the procedure. Calculations at this level of approximation are really expensive in terms of computational time and the expected gain in accuracy of the theoretical spectrum generally does not justify them. As a result, we voluntarily set aside this option for now, the procedure for calculations including "D2" being similar to HT albeit longer and more complex.

The first control sequence causes the end of the procedure if an erroneous value is given to an option and the program is unable to correct it unambiguously. The second control will be discussed in the following section and checks the validity of the data given in input. In addition to these verifications, a constant supervision of the memory usage of the procedure is performed. The latter is stopped if it has not sufficient resources to continue.

3.3 Extraction of the required data and adjustments

To perform the computations of the transition dipole moment integrals and to generate the requested UV-vis spectrum, some data are needed for both the initial and final states:

- The atomic types and coordinates
- The atomic masses
- The energy of the equilibrium structure
- The vibrational frequencies
- The nuclear displacements caused by the vibrations
- Optionally (for HT and FCHT calculations), the electronic transition dipole moment and its derivatives with respect to the final state

In this section, we will focus on the data sources and then the treatment of these data. These information are all available from GAUSSIAN, requesting the calculation of the frequencies as the type of job. However, GAUSSIAN can only treat one electronic state at a time while two are needed to simulate the spectrum of an electronic transition. Data from another source than the current calculation are needed. GAUSSIAN provides two files in output that we can use for this. The first one is the standard readable output file and the second one is the binary checkpoint file, mainly used to restart calculations.

Choosing the standard output file requires writing a specific function to read it and extract the necessary data. Two problems are inherent to this approach. Firstly, the output is written to be readable by a human. The precision is chosen to be enough for exploitation but can be insufficient for our calculations. Some additional options such as `FREQ=HPMODES` placed in the *route section* can be used to increase the number of digits printed for some values (here the normal modes), but the improvement remains limited. Secondly, depending on the type of job, the chosen method and the options set by the user, the layout of the output, as well as the displayed information, can strongly change, making very difficult the conception of a simple and versatile reading subroutine.

Nonetheless, using the GAUSSIAN output file as the source of data has several advantages. The number of files that can be read is virtually unlimited and the data can be visually controlled by directly reading the output. These characterstics make them simpler to deal with when developing new features or adding new calculation methods, especially since they can be read by other, external programs to confront the results.

The binary checkpoint file is also a really interesting source of data since values are stored inside at machine precision. Moreover, it is a file whose construction is well organized to make possible the re-use of the quantities inside it by GAUSSIAN, whatever the type of calculations required by the user when the file has been created. However, such files are more complex to handle because of their internal structure which depends partly on the processor architecture in the computer on which the calculation has been run. This problem is mostly overcome by internal tools of GAUSSIAN which are able to retrieve specific data in this file. However, these tools can only work correctly if the checkpoint file was written on a computer with the same processor architecture as the one on which GAUSSIAN is instructed to generate the spectrum[d]. To each piece of information, such as the number of atoms, the atomic coordinates, the forces and the force constants, corresponds an index. The reading tools of GAUSSIAN provide a framework which fills an array given in input with the quantities found at a given index. However, this system can only handle one checkpoint file throughout the run of GAUSSIAN.

Though the treatment of the checkpoint file appears simpler that for the GAUSSIAN output files, finding the correct indexes can be challenging. This is accentuated by the fact that only essential data are stored inside and other quantities such as the frequencies and the normal modes must be calculated from them. As a consequence, the development of the procedure was started reading external GAUSSIAN output files. While the procedure can now perfectly handle data from the current calculation and the checkpoint files, the possibility to read the output files remains.

For now, the input data can be extracted from three kinds of source: current calculation (`CALC`), checkpoint file (`CHK`), GAUSSIAN output file (`READ`). The source must be specified for both initial (`1`) and final (`2`) states. By default, data for the initial state are taken from the checkpoint file (`CHK1`) and those for the final state from the current run of GAUSSIAN (`CALC2`). If only one source is given in input, the procedure attempts to "guess" the correct source for the other state by using its default value first, and if chosen in input, the other source providing data at machine precision. Because the extraction from the GAUSSIAN output file requires a filename given in input, the procedure never assumes that the missing data source is of this kind. If the checkpoint file or the current calculation is chosen as the source for both states, the calculations end with an error. Data for both states can be chosen from GAUSSIAN output files. In this case, the module expects two filenames after the *dinautil section*.

```
NoReord FC READ1 READ2 (dinautil section)

File1
File2
```

[d] In reality, the problem is more complex, as different processor types can "produce" compatible checkpoint files. Nonetheless, it is advisable to convert the checkpoint files upon transfering them to use on different kinds of computers.

In our example, and the following examples as well, we will write the keyword `NoReord` in the *dinautil section*. This option is not directly associated to FRANCK and the generation of the spectrum. However, it has an influence on the procedure as it cancels the reordering of the frequencies and normal modes in spectroscopic order, and keep the original order by increasing energy of vibration.

Regarding the GAUSSIAN output files, if one of these file cannot be found, the procedure is stopped and an error is raised.

The sources chosen for the calculations are printed in the GAUSSIAN output file of the current job. In the case of default setting, this part of the input will be:

```
Data for initial state taken from: Checkpoint file
Data for final state taken from: Current calculation
```

In the case of information taken from GAUSSIAN output file, the filename is written.

When choosing the checkpoint file as one of the sources of data, one must be careful that the ckeckpoint file is not rewritten during the GAUSSIAN calculations. This is for example the case if geometry optimization is asked together with frequencies calculations. It is highly advisable to run geometry optimization and frequencies calculations separately when intending to perform the simulation of a UV-vis spectrum. The default procedure should be:

1. Optimization of the geometry of the initial state
2. Optimization of the geometry of the final state
3. Frequencies calculation for the initial state
4. Backup of the checkpoint file obtained before
5. Frequencies calculation for the final state + generation of the spectrum. The checkpoint file obtained at the end of the frequencies calculation of the initial state is used as the checkpoint file of this calculation (`%Chk=`*`checkpoint_file_obtained_at_point_3`*)

Once the sources of data have been chosen and are valid, the data can be extracted. Three controls are performed upon analyzing the retrieved quantities: all necessary data have been found to carry out correctly the calculations, there is no imaginary frequency (the procedure handles only stable states), and the molecular system is the same in the initial and final states. If one of these criteria is not met, the job is stopped with an appropriate error raised. In the case of FCHT or HT calculation, an additional keyword, `NSTATE=`*`value`*, allows to choose the excited state of interest. By default, it is the first excited state (`NSTATE=1`). However, it should be chosen to be the same as the state of interest chosen with the suboption `Root` of the keyword `TD`.

Next, the procedure maximizes the overall superposition of the structure in initial and final states in two steps. Such a stage is important to reduce as much as possible the spurious effects that the displacements due to translation and rotation can introduce in the spectrum. The translation can be exactly removed by superposing the center of masses of the two equilibrium structures but the rotational effects can generally only be minimized by maximizing their mutual superposition. Firstly, the structure of the initial state is reoriented along the Eckart axes after the center of

mass has been translated to the origin of the fixed spatial frame of reference. The corresponding rotation matrix is used to reorient the matrices of the Cartesian forces and Cartesian force constants. Secondly, the structure of the final state is superposed to the initial state's one and the rotation matrix is used to reorient the matrices of Cartesian forces, Cartesian force constants, and the electronic transition dipole moment and its derivatives if present. In the GAUSSIAN output file, FRANCK describes the steps of the procedure and prints the new geometries of the initial and final states.

```
Initial state structure is set in Eckart orientation.
Final state structure is superposed to it.
             New orientation in initial state
[...]
               New orientation in final state
[...]
```

The vibrational frequencies and normal modes are calculated from the matrices of forces and force constants. Once all values have been extracted and calculated when necessary, the procedure checks if modifications of some of them were requested by the user.

As a matter of fact, some values can be be corrected to improve the accuracy of the generated spectrum. These adjustments regard the difference of energy between the electronic states which can be set by the keyword `ENERINP=`*`value`* and the frequencies. There are two ways to modify the frequencies. The first one is to replace those of the initial or final state, or even both together, by new frequencies. This is done with the keywords `InFrS` followed by a digit representing the inital (`0`) or final (`1`) state. The frequencies are listed after the *dinautil section* and before the filenames of the GAUSSIAN output files if present. The lists can be written on an arbitrary number of lines and end with a blanck line. To change the vibrational frequencies of the initial and final states, one will use the keywords `InFrS0` and `InFrS1`, respectively:

```
 NoReord FC InFrS0 InFrS1

list of frequencies for the initial state

list of frequencies for the final state
```

Another method to adjust the frequencies is based on a mode-specific scaling enabled with the keyword `SclVec`. The principle is described in the subsection below.

3.3.1 A simple scheme for anharmonicity

When considering the computational strategies and the methods to calculate the Franck-Condon integrals, we limited our study to the harmonic approximations. However, improving the accuracy of simulated spectra may require to go beyond harmonicity and to take into account anharmonic effects, couplings between modes, vibrational or vibronic resonances. As hinted before, the complete treatment of the anharmonicity in the calculation of the transition dipole moment integrals requires

a complex methodology which is beyond the purpose of the present document. Nonetheless, a partial treatment of the anharmonicity could be taken into account relatively easily by using some external data. As a first step in this direction, we chose to apply anharmonic corrections to the vibrational frequencies of both ground and excited electronic states. The calculation of anharmonic frequencies still represents a complicated task, even for medium size molecular systems. Although successful approaches for vibrational state calculations have been already reported (including the treatment of molecules in electronic excited states [97, 98] and the computation of Frank-Condon factors [99]), significant problems remain, especially in relation to the dimensionality of the potential energy surface. As the size of the molecular system increases, the number of calculations needed to describe the anharmonic potential energy surface becomes so large that accurate frequency evaluations based on full-dimensional variational approaches will not be practical at least in the near future. On the other hand, in many cases the effect of anharmonicity can be accounted for in an approximate manner: thus, for example, scaling of harmonic frequencies by a (uniform or frequency dependent) factor tends to provide better agreement with experimental data. This procedure has been extensively applied for calculations of vibrational frequencies in the ground electronic state, and adequate scaling factors are proposed in the literature [100–102]; however, much less information is available concerning excited state vibrational frequencies, and the selection of a specific scaling factor would at present be rather arbitrary. Therefore, the procedure introduces a simple scheme to derive excited state mode-specific scaling factors starting from the ground state ones: in turn, the latter can be obtained e.g. by means of perturbative anharmonic frequency calculations [103], or derived by a comparison with easily accessible ground state experimental data.

To do so, three sets of frequencies are used. Harmonic frequencies of both electronic states can be calculated directly from ab initio calculations and are used as a basis. They are directly extracted from the input sources chosen by the user. The user provides a set of frequencies that will replace those of the lower state in energy (initial state in absorption and final state in emission). The reason of this choice comes from the original purpose of this method as presented previously. As a matter of fact, it is generally only possible to carry out calculations including anharmonicity for the ground state, hence the limitation. In this case, anharmonic perturbative [103] calculations are performed for the ground electronic state and the obtained frequencies are given in input after the *dinautil section* in the same way as with the keywords `InFrS0` and `InFrS1`. Because the list of frequencies is given by the user, they can be taken from a different source than anharmonic calculations, for example experimental frequencies.

After FRANCK has read the input values, it can use them to correct the frequencies in both ground and excited electronic states. In order to derive mode-specific scaling factors for each particular normal mode Q_i, the frequency scaling vector $\boldsymbol{\alpha}$ is computed first, using the formula $\alpha'_i = \vartheta'_i/\omega'_i$ where $\boldsymbol{\vartheta}'$ is the vector of the anharmonic (or experimental) frequency, and $\boldsymbol{\omega}$ is the harmonic frequency. To proceed further, we assume that, if there is a one-to-one relation between the normal modes Q'_i and Q''_i of the initial and final state, the scaling factors α'_i and α''_i are equal. However, the normal modes are in general not coincident ($\mathbf{J} \neq \mathbf{I}$), and $\boldsymbol{\alpha}'$ cannot be transferred directly to scale the frequencies of the final state. In other words, the scaling vector must be adapted to the excited state frequencies. In the case of small-amplitude vibrations, this can be obtained by expressing the normal modes of the excited state as linear combinations of the normal

coordinates of the initial state, by means of the Duschinsky transformation. The J_{ik} coefficients can now be applied to derive the relation between the initial (k) and final (i) state mode-specific anharmonicity scaling factors:

$$\alpha_i'' = \sum_k^N {J_{ik}}^2 \alpha_k' \tag{3.1}$$

An application of this procedure is described in section 4.8.

3.4 Control of the output

Before continuing with the treatment of the input data to compute the transition dipole moment integrals and generate the UV-vis spectrum, we will present in this section the possibilities of modifying the output to one's own needs or wishes. These options do not have an influence on the calculations of the overlap integrals and generally, will not increase the computational times.

A first option is available to print some of the basic matrices used throughout the calculations. The printing is enabled by the keyword `PRTMAT=`*`value`* where *`value`* is an integer formed by digits (1-7) representing one of the matrices **J**, **K**, **A**, **B**, **C**, **D**, **E**:

- `1` ← **J**: the rotation matrix, also called Duschinsky matrix. It provides information on the variations of the normal modes between the initial and final states and show their mixing.
- `2` ← **K**: the shift vector. It gives an insight on the shift of the equilibrium positions of the nuclei in the structure between the initial and final states.
- `3` ← **A**: one of the Sharp and Rosenstock matrices (see eq. 2.7). It gives in a first approximation information of the effect of the shift in frequencies on the vibrational progression for the modes of the initial state.
- `4` ← **B**: one of the Sharp and Rosenstock matrices (see eq. 2.8). It gives in a first approximation information of the effect of the shift in equilibrium positions on the vibrational progression for the modes of the initial state.
- `5` ← **C**: one of the Sharp and Rosenstock matrices (see eq. 2.9). It gives in a first approximation information of the effect of the shift in frequencies on the vibrational progression for the modes of the final state.
- `6` ← **D**: one of the Sharp and Rosenstock matrices (see eq. 2.10). It gives in a first approximation information of the effect of the shift in equilibrium positions on the vibrational progression for the modes of the final state.
- `7` ← **E**: one of the Sharp and Rosenstock matrices (see eq. 2.11), and the only one which directly accounts for the interactions between the vibrational states of the initial and final electronic states.

If one wishes to print the Duschinsky matrix **J** and the shift vector **K**, the option will be written `PRTMAT=12`. The order of the digits is unimportant and doubles are ignored. However, because of the limitations on the size of the integers in Fortran, the latter should be avoided. In our case, the

corresponding part of the output will be:

```
------------------------------------------------
   Printing matrices as requested
------------------------------------------------

 Duschinsky Matrix
 -----------------------------
 list of the elements of the matrix formatted on 5 columns

 Shift Vector
 -----------------------------
 list of the elements of the vector
```

In practice, the Duschinsky matrix written in the output file is the transpose of the matrix **J** given in equation 1.64 and represents the normal coordinates of the final state in the basis set of the normal coordinates of the initial state. In other words, the columns of the printed matrix represent the normal coordinates of the initial state (**Q**') while its rows correspond to those of the final state (**Q**").

As described in section 2.8, the prescreening method used in the procedure is independent from the bounds of the spectrum. However, to print it, the procedure needs to know its characteristics. The first two parameters are the energy of the lower bound (`SPECMIN=`*value*) and the energy of the upper bound (`SPECMAX=`*value*). Their values, given in inverse centimeters, can be defined relatively to the energy of the transition between the fundamental vibrational states of both electronic states or absolutely. By default, the energies are relative and the spectrum ranges from `SPECMIN=-1000` cm^{-1} to `SPECMAX=+8000` cm^{-1} for an absorption spectrum and from `SPECMIN=-8000` cm^{-1} to `SPECMAX=+1000` cm^{-1} in the case of emission. Switching to absolute values is done by inserting the keyword `NORELI00` in the *dinautil section*. When doing so, both bounds are expected to be in absolute energies. Hence, it is highly advisable to redefine both `SPECMIN` and `SPECMAX` in this case. A control will check that the lower bound of the spectrum has a lower energy than the upper bound (`SPECMAX` > `SPECMIN`). If not, the procedure stops and an error is raised.

Another important characteristic of the spectrum is the interval of energy separating two points of measurement. The lower is this interval, the higher is the resolution of the spectrum. By default, the value, defined with the keyword `SPECRES`, is set to `8` cm^{-1}. It is important to note that a really high resolution can greatly hamper the speed of the calculation, especially in association with the full-width at half-maximum used for the convolution described below.

In reality, we never obtain stick absorption or emission spectra but continuous ones comprised of bands. The reasons of the broadening of the spectral lines over a range of frequencies instead of a single frequencies are generally manifolds. A first, natural broadening is a consequence of the uncertainty principle which states that the shorter is the lifetime Δt of a system in an excited state, the higher is the imprecision in the energy ΔE of this state, following the relation $\Delta E \Delta t \geq \hbar/2$. As a consequence, the states involved in a transition whill have slightly different energies in each

molecule. Such a broadening is homogeneous and the distribution is represented by a Lorentzian distribution. Two other effects are due to the conditions of the system when the analysis are carried out. When exposed to temperature, each molecule is in motion if the thermal energy is sufficient. Molecules traveling toward the detector of the spectometer will have transition frequencies which differ from those at rest or moving away. The higher is the temperature, the larger is the distribution of velocities, causing an additional broadening of the spectral bands. This phenomenon is especially present in liquid phase and leads to an inhomogeneous broadening represented with a normal distribution (Gaussian function) In this case, the presence of nearby particles also cause interactions that will modify even more the energies of transition for each system, leading to an homogeneous distribution.

To account for these effects, the theoretical spectrum can be convoluted with a Gaussian-type exponential or a Lorentzian function. The Half-Width at Half-Maximum can be set with the keyword `SPECHWHM`. By default, it is `135` cm^{-1}. It should be noted that the values for these four options, `SPECMIN`, `SPECMAX`, `SPECRES` and `SPECHWHM` are not directly summarized in the output of FRANCK, which means that only non-default parameters are printed when the *dinautil section* is rewritten (see section 3.2 for more details). The resolution should be chosen with care in relation with the desired convolution of the spectrum. A high resolution with a large width of bands can be extremely time-consuming.

In our procedure, we assumed that the inhomogeneous breadening is dominant, so the convolution is obtained with a normal distribution. However, a Lorentzian function will be added in the future to simulate a natural broadening.

The probability density is represented by a Gaussian function f_G :

$$f_G(\omega) = \frac{1}{\varsigma\sqrt{(2\pi)}} \exp\Big(-\frac{(\omega-\omega_0)^2}{2\varsigma^2}\Big) \tag{3.2}$$

where ω_0 is the expected value, in our case the exact energy of the transition, and ς is the standard deviation. The standard deviation is related to the half-width at half-maximum with the expression:

$$\varsigma = \frac{\texttt{SPECHWHM}}{\sqrt{2\ln(2)}}$$

Replacing ς by `SPECHWHM`, the Gaussian function can be written:

$$f_G(\omega) = \frac{\ln(2)}{\texttt{SPECHWHM}\sqrt{\pi}} \exp\Big(-\frac{\ln(2)(\omega-\omega_0)^2}{\texttt{SPECHWHM}^2}\Big) \tag{3.3}$$

About 99.7% of the surface is covered within an interval of 6 standard deviations about the center of the peak ω_0 ($\omega = 3\varsigma$)[e]. This corresponds to a range of about $5.1 \times$ `SPECHWHM`. A safer zone covering $3.5 \times$ `SPECHWHM` on each "side" of the expected value is chosen to represent our Gaussian function. In this case, the number of intervals (*Linewidth*) on which will span a single band is given by the formula:

$$Linewidth = 7 * \left\lfloor \frac{\texttt{SPECHWHM}}{\texttt{SPECRES}} \right\rfloor$$

where $\lfloor x \rfloor$ represents the floor function.

[e]This value corresponds to the confidence interval and is given by multiple tables available in literature.

For each transition, the intensity of the line of absorption or emission, $\sigma_{\boldsymbol{v}',\boldsymbol{v}''}$ is calculated thanks to the formula given in equation 1.20 or 1.21, respectively:

$$\text{Absorption} : \sigma_{\boldsymbol{v}',\boldsymbol{v}''} = \frac{4\pi^2\omega}{3c}\rho_{v'}\left|\langle\, \boldsymbol{v}' \mid \boldsymbol{\mu}_{if} \mid \boldsymbol{v}'' \,\rangle\right|^2\delta(E_{\boldsymbol{v}''} - E_{\boldsymbol{v}'} - \hbar\omega_0)$$

$$\text{Emission} : \sigma_{\boldsymbol{v}',\boldsymbol{v}''} = \frac{4\omega^3}{3\hbar c^3}\rho_{v'}\left|\langle\, \boldsymbol{v}' \mid \boldsymbol{\mu}_{if} \mid \boldsymbol{v}'' \,\rangle\right|^2\delta(E_{\boldsymbol{v}''} - E_{\boldsymbol{v}'} + \hbar\omega_0)$$

where $\rho_{v'}$ is the Boltzmann population of the vibrational initial state, $E_{\boldsymbol{v}'}$ and $E_{\boldsymbol{v}''}$ are the energies of the initial and final vibronic states, respectively, and ω_0 is the energy of the transition.

As written before, the calculated spectrum is not continuous but discretized in a finite number of measuring points $NPoints$ equal to:

$$NPoint = \left\lceil \frac{\texttt{SPECMAX} - \texttt{SPECMIN}}{\texttt{SPECRES}} \right\rceil$$

where $\lceil x \rceil$ represents the ceil function.

While the center of the band remains the energy of the transition ω_0, in practice, the center on the discretized spectrum is chosen to be the higher measuring point with an energy lower than ω_0. Starting from this point, the band is sampled on the $Linewidth/2$ measuring point before and after the central point. For each sampling point i, the intensity corresponding to the convolution $\sigma^{\mathrm{c}}_{\boldsymbol{v}',\boldsymbol{v}''}$ is calculated with the relation:

$$\sigma^{\mathrm{c}}_{\boldsymbol{v}',\boldsymbol{v}''} = \sigma_{\boldsymbol{v}',\boldsymbol{v}''} \times \frac{\ln(2)}{\texttt{SPECHWHM}\sqrt{\pi}} \exp\left[\ln(2)\left(\frac{\omega_i - \omega_0}{\texttt{SPECHWHM}}\right)^2\right]$$

where ω_i is the energy of the measuring point i.

Because of the sampling to compute the spectral bands, a small interval of energy between measuring points associated to a large half-width at half-maximum will cause a large number of calculations for each transition. Nonetheless, a resolution too low can slightly distort the accuracty of the final spectrum.

It is noteworthy that a different spectrum is calculated for each class $\mathcal{C}_n$ as well as the transition between the fundamental vibrational states, treated as a single class $\mathcal{C}_0$. The spectra are summed to give the overall spectrum:

```
 +------------------+
 |  Final Spectrum  |
 +------------------+

Axis X = Energy (in cm^-1)
Axis Y = Intensity

 ------------------------------------------------------------
 Axis_X  Axis_Y
```

It is possible to request that each spectrum is printed with the keyword `ALLSPECTRA`. A new section following the general spectrum is created in the output file.

```
+--------------------+
|  Specific Spectra  |
+--------------------+

Number of spectra: xx
-- Spectrum num.  1: I00 spectrum --
-- Spectrum num.  2: Overtones spectrum --
-- Spectrum num. yy: yy-states combinations spectrum --
-- Previous line is repeated until the last class considered
      in the calculations

------------------------------------------------------------
Axis_X Axis_Y
```

xx is the number of spectra, that is to say, the highest class + 1. All the spectra are listed one after another with an horizontal dashed line to separate each one.

Only the convoluted spectra are printed. Two ways are possible to obtain the stick spectrum. The first one is to choose `SPECHWHM` in order for *Linewidth* to be equal to 0. It is enough to define it lower than `SPECRES`. Another method is to generate it from the assignment printed in the output file.

Indeed, to facilitate the analysis of the UV-vis spectra, the procedure prints the transitions whose probability is above a given threshold. By default, its value is set to 1% of the squared transition dipole moment integral $\langle\, \mathbf{0'} \mid \boldsymbol{\mu}_{if} \mid \mathbf{0''} \,\rangle$. It can be changed with the keyword `PRTINT=`*value* where *value* is used as a factor of $\left|\langle\, \mathbf{0'} \mid \boldsymbol{\mu}_{if} \mid \mathbf{0''} \,\rangle\right|^2$ to define the minimum probability a transition must have to be shown. The list of all transitions satisfying this condition is given in a specific section introduced by the title "`Overlap integrals computation`" with the following structure. Firstly, the energy corresponding to the transition between the fundamental vibrational states is printed in inverse centimeters. This energy is used as the reference, the transition energies being relative to it. Then, the listing of the transitions is divided into sections corresponding to the classes ordered by increasing number of excited oscillators, starting to 0. For each transition, five pieces of information are given: the non-zero quantum numbers of the initial state $\boldsymbol{v'}$, the non-zero quantum numbers of the final state $\boldsymbol{v''}$, the relative energy of the transition $\omega_0 - (E_{\mathbf{0''}} - E_{\mathbf{0'}})$, the squared transition dipole moment integral $\left|\langle\, \boldsymbol{v'} \mid \boldsymbol{\mu}_{if} \mid \boldsymbol{v''} \,\rangle\right|^2$ and the intensity of the spectrum line $\sigma_{\boldsymbol{v'},\boldsymbol{v''}}$.

```
 ------------------------------------------------
    Overlap integrals computation
 ------------------------------------------------

 Energy of the 0-0 transition:  E(0,0) cm^(-1)

 Notes about the overlap integrals description:
  - DeltaE is the relative energy of the transition (wrt 0-0)
  - TDMI**2 is the squared transition dipole moment integral
  - Intensity is the line intensity (absolute value)

[...]
 -- To: single overtones --

Initial State: <list of the excited modes i: i^v |
Final State: |list of the excited modes i: i^v >
     DeltaE =  value  | TDMI**2 = value , Intensity = value
   ......................................
```

By adding the intensities given for each transition with respect to the energy of the latter, it is possible to compute the stick spectrum. Only the transition with a low intensity are discarded, so that the resulting stick spectrum should have a satisfactory accuracy. More transitions are obviously obtained by lowering the threshold using `PRTINT`. However, the printing of the information is done directly after the transition dipole moment integral has been evaluated. This permits to avoid increasing the storage by saving in memory the intensities to write them later in the output. The drawback of this choice is that, the lower is the threshold, the more interaction the procedure will have with the output. This time-consuming operation can noticeably slow down the calculations so it is advisable to avoid setting a very low threshold. Additionnally, it is possible to use a high threshold to lighten the output when the aim is to assign the most intense transitions only. It is important to note that in the case of weakly-allowed or dipole-forbidden transitions, the intensity of the reference transition can be really low. Consequently, the default parameter can give a very long list of transitions with very low probabilities.

3.5 Calculation of the required matrices

Going back to the calculations, the procedure, after the extraction and recalculation of the data, follows a common trunk at first. This part is centered on the construction of the generic matrices which will be mostly used in the recursion formulae to compute the overlap integrals. The transformation matrices $\mathbf{L}$ are computed from the nuclear displacements caused by the vibrations. As discussed before, an important property of these matrices is the orthonormalization of their columns, to that we have the relation:

$$\sum_{i=1}^{N_A} \sum_{\tau=x,y,z} {\mathrm{L}_{i\tau j}}^2 = 1$$

To meet this requirement, the elements of $\mathbf{L}$ are calculated with the following formula:

$$\mathrm{L}_{i\tau j} = \frac{\sqrt{m_i}\,\delta_{n\,i\tau j}}{\sum_{k=1}^{N_A}\sum_{\varrho=x,y,z} m_k\left(\delta_{n\,k\varrho j}\right)^2} \tag{3.4}$$

where $\delta_{n\,i\tau j}$ represents the displacement of the nucleus i along one Cartesian coordinate τ induced by the vibration j. m_i is the mass of atom i. Practically, i and τ are accounted for in a single column of dimension $3 \times N_A$.

Once $\mathbf{L}'$ and $\mathbf{L}''$ have been calculated, the Duschinsky matrix is obtained using the relation $\mathbf{J} = \left(\mathbf{L}'\right)^{\mathrm{T}}\mathbf{L}''$. It is possible to ignore the mode coupling so that the rotation matrix is equal to identity with the keyword `JIdent`. The principle, however, is different from the scheme used by Ervin *et al.* [81, 87] in their parallel mode approximation. The model Duschinsky matrix in this case is used to simulate a spectrum when the rotation of the normal modes is purely not taken into account. Because of the difficulty encountered by Ervin *et al.* in their approximation of the $\mathbf{J}$ with several systems, a simpler, but less accurate, procedure was devised here. Its primary objective was to simulate the behavior of simple models such as the Linear Coupling Method described more precisely in the Introduction of this document, which was used as a reference in section 4.8. Hence, this approximation of the Duschinsky matrix can be used to compare and gain a better understanding on the consequences of ignoring, or inversely taking into account, some characteristics of the potential energy surface of the excited state.

$\mathbf{K}$ is then calculated from the relation $\mathbf{K} = \left(\mathbf{L}'\right)^{\mathrm{T}}\mathbf{M}^{1/2}\mathbf{S}$ where $\mathbf{S}$ is the difference of equilibrium geometries in Cartesian coordinates between the initial and final state ($\mathbf{S} = \mathbf{X}'_{\mathrm{eq}} - \mathbf{X}''_{\mathrm{eq}}$). The calculations of the transition dipole moment integrals are all performed in atomic units, which require some conversions from the units used by GAUSSIAN (such as the atomic mass unit).

The electronic transition dipole moment and its derivatives are then evaluated depending on the *type of calculation* requested by the user. In case of FC calculations, the zeroth-order of the Taylor series of $\boldsymbol{\mu}_{if}$ with respect to the normal coordinates of the final state given in equation 1.29 can be considered as a constant scaling factor for the line intensities in the spectrum. As a consequence, its exact value is ignored and the norm of $\boldsymbol{\mu}_{if}(\mathbf{Q}''_0)$ is currently set to unity.

When dealing with an approximation beyond the Franck-Condon principle, the procedure requires the electronic transition dipole moment and its derivatives with respect to the normal modes. The latter are not directly available from GAUSSIAN which gives the derivatives with respect to the displacements in Cartesian coordinates, $\partial\boldsymbol{\mu}_{if}/\partial\mathbf{X}''$. The relation between these derivatives and those with respect to the normal coordinates is straightforward, using the transformation given in equation 1.28:

$$\left(\frac{\partial\mu_{if}(\tau)}{\partial Q''_k}\right)_0 = \sum_{l=1}^{N_A}\sum_{\varrho=x,y,z} \mathrm{L}_{lk\varrho}\mathrm{M}_{kk}^{-1/2}\left(\frac{\partial\mu_{if}(\tau)}{\partial X''_{l\varrho}}\right)_0 \tag{3.5}$$

where M_{kk} is an element of the diagonal matrix $\mathbf{M}$ which is equivalent to the mass m_k used in equation 3.4.

Currently, the procedure only deals with the Taylor expansion of the electronic transition dipole

moment about the equilibrium geometry of the final state. As a result of this choice, it does not expect this information in the data of the initial state. For example, if the calculation of an emission spectrum from an electronic excited state to the ground state is requested by the user, it will look for the derivatives of the electronic transition dipole moment in the source file of the latter. If they have only been calculated with respect to the excited state, then the procedure will be unable to generate the FCHT/HT spectrum. Such an issue can be resolved by calculating the derivatives also in the ground state but this slightly reduces the desired flexibility of our procedure. It should be noted that this has not been a problem for our studies presented in chapter 4. Nonetheless, we describe in the conclusion of this chapter a scheme currently in development to overcome this limitation by allowing the procedure to extract from any data sources and carry out correctly the calculations.

It should be noted that depending on the method chosen in the *route section*, the derivatives of the electronic transition dipole moment might not be calculated. This behavior can be changed by switching from analytic frequencies to numerical frequencies. When choosing to extract the input data from the GAUSSIAN output files, one should be careful that all necessary data to evaluate the derivatives of $\boldsymbol{\mu}_{if}$ are printed. This problem can happen when the molecule has a symmetry different from C_1. In this case, only a part of the derivatives will be written, the other ones being infered thanks to the symmetry. However, FRANCK is currently not able to do so. Consequently, when planning to use the output file as input data for the simulation of spectra, it is advisable to disable the treatment of the symmetry with the option `NoSymm` in the *route section*. Unfortunately, the drawback of this setting is an increase of the computational costs.

In the GAUSSIAN output file, the values of $\partial\boldsymbol{\mu}_{if}/\partial\mathbf{X}''$ are not directly available. They are recovered by finite differences, the electronic transition dipole moment being computed after each displacement of an atom at a time along a single Cartesian coordinate. Practically, $\boldsymbol{\mu}_{if}(\mathbf{X}''_{\text{eq}})$ is calculated first and printed. Then, the first atom is slightly displaced along the x axis by a quantity $-\delta_{\text{step}}$ and the corresponding $\boldsymbol{\mu}_{if}(X''_{1x-})$ is calculated. The same atom is then displaced from its equilibrium position along the same axis, this time by a quantity $+\delta_{\text{step}}$ and $\boldsymbol{\mu}_{if}(X''_{1x+})$ is calculated. The same procedure is done along the axes y and z and for all atoms of the system. The derivatives of $\boldsymbol{\mu}_{if}$ with respect to the Cartesian coordinates are then obtained with the relation:

$$\left(\frac{\partial\mu_{if}(\tau)}{\partial X''_{k\tau}}\right)_0 = \frac{\boldsymbol{\mu}_{if}(X''_{k\tau+}) - \boldsymbol{\mu}_{if}(X''_{k\tau-})}{2\times\delta_{\text{step}}} \tag{3.6}$$

δ_{step} is a constant set by default at $10^{-3}\mathring{A}$. When evaluating the first derivatives, FRANCK assumes that δ_{step} has this value.

Unfortunately, standard calculations do not provide a way to compute easily the second derivatives of the electronic transition dipole moment. Since only one atom is moved at a time, the available data can only permit to evaluate the diagonal elements of these derivatives. Hence, in the current implementation of the procedure, the terms of the second-order of the Taylor series of the electronic transition dipole moment, and so the calculations at the level of approximation "D2", are restricted to the diagonal terms. The evaluation of these derivatives is possible by supposing that the first derivatives are obtained through finite differences by displacing the atoms of a distance $\delta_{\text{step}}/2$ instead of δ_{step}, leading to the coordinates $X''_{k\tau(+/2)}$ or $X''_{k\tau(-/2)}$ instead of $X''_{k\tau+}$ or $X''_{k\tau-}$,

respectively. Applying twice the finite differences, one can write:

$$\begin{aligned}\left(\frac{\partial^2\mu_{if}(\tau)}{\partial {X''_{k\tau}}^2}\right)_0 &= \frac{\partial}{\partial X''_{k\tau}}\left(\frac{\partial\mu_{if}(\tau)}{\partial X''_{k\tau}}\right)_0 = \frac{\partial}{\partial X''_{k\tau}}\left(\frac{\boldsymbol{\mu}_{if}(X''_{k\tau(+/2)}) - \boldsymbol{\mu}_{if}(X''_{k\tau(-/2)})}{\delta_{\text{step}}}\right)\\ &= \frac{\boldsymbol{\mu}_{if}(X''_{k\tau+}) - \boldsymbol{\mu}_{if}(X''_{\text{eq}}) - (\boldsymbol{\mu}_{if}(X''_{\text{eq}}) - \boldsymbol{\mu}_{if}(X''_{k\tau-}))}{{\delta_{\text{step}}}^2}\\ &= \frac{\boldsymbol{\mu}_{if}(X''_{k\tau+}) + \boldsymbol{\mu}_{if}(X''_{k\tau-}) - 2\boldsymbol{\mu}_{if}(X''_{\text{eq}})}{{\delta_{\text{step}}}^2}\end{aligned} \tag{3.7}$$

While being partial, accounting for the diagonal terms of the second derivatives can still provide a refinement of the accuracy of the generated spectrum, particularly in the case of symmetric molecules. This also sets a basic framework for a generalization taking into account all elements of the second derivatives. However, the computational costs at this level of approximation of the electronic transition dipole moment are quite high so that the full treatment of the second derivatives should be reserved to very specific cases.

The rest of the common trunk compute the Sharp and Rosenstock matrices **A**, **B**, **C**, **D** and **E** and finally the overlap integral between the fundamental vibrational states of the initial and final states $\langle\, \mathbf{0'} \mid \mathbf{0''} \,\rangle$. This part is a direct implementation of the formulae given previously.

The treatment of the transitions is not directly done by the second central subroutine `FRANCK2` but by specific subroutines (that can be generically called `FRANCK3`) which are called depending on the *type of calculation* as shown in figure 3.1.

3.6 General structure to organize the computation of the transition dipole moment integrals

As hinted before, FRANCK has several, similar subroutines to manage the calculations of the transition dipole moment integrals depending on the highest order considered in the Taylor series given in equation 1.29. Indeed, the higher the order, the more calculations are necessary to evaluate the intensity of a single transition. Without even considering the recursion, equation A.1 shows that only one overlap integral is needed for FC calculations, while $2N$ are required for HT calculations and $4N^2$ for D2. While, in the current procedure, D2 is limited to the diagonal terms, theoretically lowering its order of magnitude to N, the recursion formulae for some terms such as $\langle\, \boldsymbol{v'} \mid \boldsymbol{v''} + 2''_k \,\rangle$ are more cumbersome than for HT calculations. Based on this observation, it is obvious that to be as efficient as possible, our calculations must be restricted to the necessary terms. Different ways are available to tackle the matter such as using conditional blocks inside the calculations. However, there are several disadvantages to this approach. The first one is that the tests must generally be performed for each transition. For a huge number of transitions, as it is the case in most simulations of UV-vis spectra, this can noticeably hamper the efficiency of the computations. A second problem is related to the difficulty to optimize the calculation schemes because each "order" must be treated separately, thus preventing the contemporaneous treatment of similar overlap integrals from different orders. Finally, the memory requirements are not identical for each case and an optimum management will demand additional tests. The direct consequence for the code is a risk

of overload, reducing its readibility. This makes it more difficult to maintain it or extend it.

In our implementation inside GAUSSIAN, the choice was made to avoid this kind of tests and to design specific routes depending on the highest term of the electronic transition dipole moment Taylor series, FC, HT or D2. The structure of each branch is very similar and is shown schematically in figure 3.2. An obvious drawback of this method lies in the similarity of these structures and the redundance they create in the program source. This is unavoidable in this case and has a consequence on the easiness to maintain and develop the code as all branches must be treated together while retaining the particularities of the calculations, such as a slightly different memory usage and construction. These specificities will be discussed in greater details in section 3.7. From experience, be it the overall performances of the procedure or the clarity of the code, the inconvenience described previously is largely counterweighted by the possibilities offered by the division of the treatment of the transition dipole moment integrals in parallel segments. Furthermore, this partition also allows a tailoring of the calculations with respect to the levels of approximation or also devise specific schemes to increase the efficiency for a particular branch.

The general procedure is as follows in each branch. Firstly, the transition dipole moment integral between the fundamental vibrational states, $\langle\ \mathbf{0}'\ |\ \boldsymbol{\mu}_{if}\ |\ \mathbf{0}''\ \rangle$, is calculated and the convoluted spectrum is generated (spectrum $\mathcal{C}_0$). Then, the subroutines for the calculations of the classes $\mathcal{C}_1$ and $\mathcal{C}_2$ are called in succession. Finally, a generic subroutine is called to treat separately each class $\mathcal{C}_n$ until a maximum n is reached. This limit can be chosen by the user with the parameter `MAXBANDS=value` or decided by the program upon some criteria. By default, the highest class is $\mathcal{C}_7$ but it might happen than going up to this limit or the one set by the user is impossible or pointless. Automatically, the program will also lower `MAXBANDS` to N if it is accidentally given higher than the number of normal modes. While performing the prescreening, for a given class $\mathcal{C}_n$, the procedure can obtain an insufficient number of modes with a non-zero maximum number of quanta $\boldsymbol{v}''_{\max}$. In this case, it stops the calculations and does not try the higher classes. This problem is illustrated with the molecule SF_6^- presented in section 4.3. Finally, it is also possible to set a minimum improvement in the spectrum progression that must be achieved between two classes $\mathcal{C}_n$ and $\mathcal{C}_{n+1}$. The spectrum progression (SP) represents the fraction of the analytic limit I^{tot} defined in section 2.9 which has been recovered summing all the squared transition dipole moment integrals weighted by the Boltzmann population of the initial state calculated up to now. Since the Taylor series of the electronic transition is about the equilibrium geometry of the final

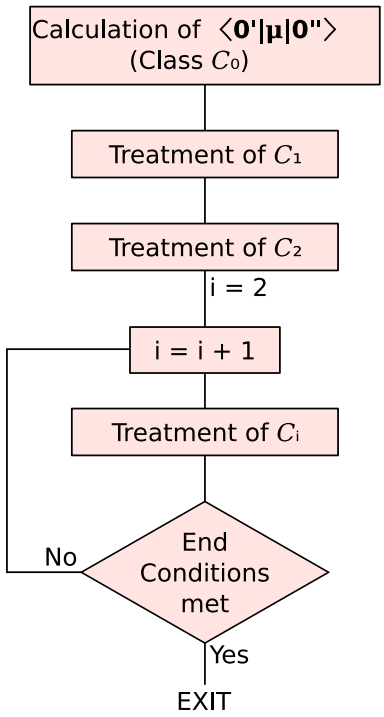


Figure 3.2: General diagram of a branch. For each class, a specific spectrum is generated and the overall spectrum progression is printed at the termination of the class.

state in our calculation, we use $I^{\mathrm{tot}}_{\mathbf{Q}''}$ as the reference analytic limit:

$$SP = \frac{\sum_i \rho_i \sum_f \sum_{\tau=x,y,z} \langle\, \boldsymbol{v}_i \mid \mu_{if}(\tau) \mid \boldsymbol{v}_f \,\rangle^2}{I^{\mathrm{tot}}_{\mathbf{Q}''}} \tag{3.8}$$

To keep a consistent notation with respect to the summation indexes, we used the notations $\boldsymbol{v}_i$ instead of $\boldsymbol{v}'$ and $\boldsymbol{v}_f$ instead of $\boldsymbol{v}''$.

The spectrum progression is evaluated after the completion of each class (including $\mathcal{C}_0$) and printed in the output:

```
 --------------------
++ Spectrum progression:  percentage of completion
```

This result is written in the section dedicated to the assignment, right after the list of the most probable transitions and just before the line indicating the treatment of the following class.

By default, there is no minimum spectrum progression required between two following classes. A different value can be set with the parameter `SPDelta=`*`value`*. If `SPDelta` is greater than 0, the calculations can be terminated if one of the three conditions is met: the progression between two classes ($SP_n - SP_{n-1}$) is lower than the threshold `SPDelta`, the highest class $\mathcal{C}_{\texttt{MAXBANDS}}$ has been handled, or an insufficient number of non-zero maximum quantum number has been chosen by the prescreening method. It is noteworthy that the current implementation of the method does not anticipate the value "$SP_n - SP_{n-1}$". Consequently, the gain of spectrum progression in a given class $\mathcal{C}_n$ is only known after it has been entirely managed, which means that if `SPDelta` is the first condition to be met, the procedure will stop after the class $\mathcal{C}_n$ for which $SP_n - SP_{n-1} <$ `SPDelta`. If the computations end for one of the reasons given above before class `MAXBANDS` has been done, FRANCK indicates the cause of the termination in the output and ends. In that case, it does not raise an error but returns the number of the last class it has completely treated. This value is used by the main subroutine to print the specific spectra required through the option `ALLSPECTRA`.

It is important to note that except $\mathcal{C}_1$ and $\mathcal{C}_2$, necessary for the prescreening methods, the other classes are independent from each other. An interesting possibility is to treat them in parallel to speed up the generation of the spectra. Technically, however, this parallelization brings up several difficulties ranging from minor to complex ones. A first, noticeable one is a more untidy output at the level of the transition assignment. Indeed, the printing of the information about the most likely transitions follows immediately the calculation of the corresponding transition dipole moment integral. Consequently, it is no more possible to partition them by class such as in a serial approach. A possible solution is to print the class with each assignment so that the output can be simply post-processed with a simple script. More serious matters concern the handling of the procedures to force the early termination of the calculations and the memory management which becomes more difficult, the resources needs automatically increasing. The control procedures mean here the threshold on the spectrum progression and the insufficient number of quanta to perform the calculations. The latter is not really cumbersome as the treatment of the class is terminated immediately after the evaluation of the maximum quantum numbers has been performed. Let us suppose that all classes $\mathcal{C}_n$ with n ranging from 3 to `MAXBANDS` are dealt contemporaneously and

the prescreening finds a number of excited modes N_{exc} lower than `MAXBANDS`. Then all the classes above N_{exc} are done in a short time, the prescreening being far less expensive in computational time than the actual calculations of the overlap integrals. If the number of available processors is lower than "`MAXBANDS - 2`" then the classes higher than N_{exc} that should have been treated later once some processors have been freed do not even need to be considered. The matter of the spectrum convergence is more complex because of the simultaneous handling of the classes. It is straightforward to see that the number of combinations of the excited modes, represented by the binomial coefficient ${}_NC_n$, increases with the class until it reaches a maximum when half of the normal modes are excited. In practice, it is pointless for a vast majority of systems to go beyond the class $\mathcal{C}_{10}$. This means that for a molecule with more than 9 atomes, the number of combinations, and so the computational costs, will keep increasing with the class. If GAUSSIAN is run in parallel on similar processors, one should expect class $\mathcal{C}_3$ to finish before $\mathcal{C}_4$ and so on until the last class. A simple strategy would be to evaluate the spectrum progression from the previous class once the treatment of a class $\mathcal{C}_n$ has ended ($\Delta = SP_n - SP_{n-1}$). In this case, if the progression relative to the class $\mathcal{C}_n$ is lower than `SPDelta`, the calculations are stopped at this point for the higher classes, preventing them to continue until their normal termination. While it can technically be done, this creates some undesirable side-effects. For example, if transitions of "aborted" classes have been printed in the section of output dealing with the assignment, the corresponding spectrum should be printed. However, the latter is incomplete, which makes its analysis rather approximative. Consequently, it is preferable not to print them to avoid misinterpretations between two calculations when the treatment of the same class has been stopped at different moments or one of them has been terminated correctly in only one case. Currently, the safer choice is to deactivate `SPDelta` when a parallel calculation is run.

It is also obvious that a storage array is necessary for each class to save the computed overlap integrals used in the recursion formulae. A direct consequence is an increase in memory requirement when the parallelization run. However, as hinted at the end of sections 2.6 and 2.8, we restrict the storage to small subsets so that the storage really needed for a class is indeed very limited. The creation of the subsets will be discussed more in-depth in the following section.

3.7 Management of the classes and calculation of the transition dipole moment integrals

As mentioned above, the classes are treated in three groups, $\mathcal{C}_1$, $\mathcal{C}_2$ and $\mathcal{C}_n$ which corresponds to all the classes starting from $\mathcal{C}_3$. This partition follows the principle of the prescreening method described in section 2.8 and discussed in the following section. Here, we will suppose that the limit $\boldsymbol{v}''_{\text{max}}$ is known and will focus on the calculation of the transition dipole moment integrals. While the partition in three kinds of classes follows the general scheme of the evaluation method implemented in FRANCK, the structure to compute the spectral lines is designed to be generic, so that implementing a new method or modifying this one can be done with little consequences on the calculation procedure.

In each class, the transitions are treated by groups corresponding to a specific combination of excited modes of the final state. In the case of $\mathcal{C}_1$, we have N *sets*, one for each mode of the final

state. For $\mathcal{C}_2$, there are $N \times (N-1)$ *sets* to represent the combinations of 2 modes of the final state. For higher classes, the problem is similar but the number of combinations depend on the non-zero maximum quantum numbers evaluated through the prescreening. This number N_C is given by the relation:

$$N_C = \frac{\underline{N}!}{(\underline{N}-n)!n!}$$

where n represents the class and $\underline{N}$ the number of non-zero maximum quantum numbers evaluated through the prescreening.

Each *set* is treated separately and independently. Hence, the storage array for the overlap integrals can be reset after each *set* has been terminated. In the most simple form, the *sets* can be listed with the use of nested loops taking care to avoid double counts. In the following discussion, we will restrict our study to the transitions from the fundamental vibrational state of the initial state, discarding in this way the thermal effects. Now, let us consider a generic class $\mathcal{C}_n$ and some combination of n excited modes, thus choosing a single *set* of the class. These selected modes will generate a subset $\underline{\boldsymbol{v}}''_{\max}$ of dimension n, containing only the non-zero maximum quantum numbers of the final state. We also need to define a second set of dimension n, $\underline{\boldsymbol{v}}''_{\mathrm{id}}$, which will store the proper index of each excited mode. For example, if we consider the first mode of the final state, j, taken in the *set*, its number of quanta is given by the subset $\underline{v}''_1$ and its order in the entire list of the normal modes of the final state can be retrieved inside the *set* by $\underline{v}''_{1_{\mathrm{id}}}$, so that we can write the equivalence $v''(\underline{v}''_{1_{\mathrm{id}}}) = \underline{v}''_j$.[f] The number of transitions to compute in the *set* is the product of the n maximum quantum numbers to take into account for each mode i of the *set*, $\underline{v}''_{i_{\max}}$. To evaluate the intensity of one of these transitions, the equation of the transition dipole moment integral given in equation A.1 is used, taking into account all the terms needed for the *type of calculation* requested by the user. We will study here the most common *types of calculation*, starting by FC and then the case of FCHT and HT, which are in practice equivalent in term of computational costs.

In the case of FC the transition dipole moment integral is

$$\langle\, \mathbf{0}' \mid \boldsymbol{\mu}_{if} \mid \boldsymbol{v}'' \,\rangle = \langle\, \mathbf{0}' \mid \boldsymbol{v}'' \,\rangle \tag{3.9}$$

We assumed here $|\boldsymbol{\mu}_{if}(\mathbf{Q}''_0)| = 1$ as done by our procedure. Using the recursion formula given in equation 2.31, equation 3.9 can be written:

$$\begin{aligned}\langle\, \mathbf{0}' \mid \boldsymbol{\mu}_{if} \mid \boldsymbol{v}'' \,\rangle = \frac{1}{\sqrt{2v''_k}}\Big[&\mathrm{D}_k\langle\, \mathbf{0}' \mid \boldsymbol{v}'' - 1''_k \,\rangle + \sqrt{2(v''_k-1)}\,\mathrm{C}_{kk}\langle\, \mathbf{0}' \mid \boldsymbol{v}'' - 2''_k \,\rangle \\ &+ \sum_{\substack{l=1 \\ l\neq k}}^{N} \sqrt{2v''_l}\,\mathrm{C}_{kl}\langle\, \mathbf{0}' \mid \boldsymbol{v}'' - 1''_k - 1''_l \,\rangle\Big]\end{aligned} \tag{3.10}$$

We presented here the general case but our *set* is limited to n non-zero quantum numbers at most. It is straightforward to see that the summation will only run on the modes of the *set*. We define $\underline{\boldsymbol{v}}''$ a subset of $\boldsymbol{v}''$ containing only the modes selected in the *set*. Additionally, the recursion can be carried out with any mode i as the starting differentiation (see equation 2.24), so that it can be chosen equal to the last mode $\underline{n}$ of the *set* with a non-zero quantum number for simplicity.

[f] $\underline{v}''_{1_{\mathrm{id}}}$ in $v''(\underline{v}''_{1_{\mathrm{id}}})$ represents an index of $\boldsymbol{v}''$.This notation was chosen here to avoid a pile-up of subscripts

Finally, it can be pointed out that all overlap integrals required in equation 3.10 are relative to $|\, \underline{\boldsymbol{v}}'' - 1''_k \,\rangle$. We denote the latter with a circle above $\boldsymbol{v}$ so that $|\, \underline{\mathring{\boldsymbol{v}}}'' \,\rangle = |\, \underline{\boldsymbol{v}}'' - 1''_k \,\rangle$ Consequently, equation 3.10 can be written:

$$\langle\, \mathbf{0}' \mid \boldsymbol{\mu}_{if} \mid \boldsymbol{v}'' \,\rangle = \frac{1}{\sqrt{2\underline{v}''_{\underline{n}}}} \Big[\mathrm{D}(\underline{v}''_{\underline{n}_{\mathrm{id}}}) \langle\, \mathbf{0}' \mid \underline{\mathring{\boldsymbol{v}}}'' \,\rangle + \sum_{k=1}^{n} \sqrt{2\mathring{\underline{v}}''_k}\, \mathrm{C}(\underline{v}''_{\underline{n}_{\mathrm{id}}}, \underline{v}''_{k_{\mathrm{id}}}) \langle\, \mathbf{0}' \mid \underline{\mathring{\boldsymbol{v}}}'' - 1''_k \,\rangle \Big] \tag{3.11}$$

It is noteworthy that only portions of the vector $\mathbf{D}$ and the matrix $\mathbf{C}$ are necessary to carry out the calculations in a *set*. While the equation corresponds to a single row of $\mathbf{C}$, one should keep in mind that our strategy of a partial storage of the overlap integrals implies that at the beginning of a *set*, only the overlap integral $\langle\, \mathbf{0}' \mid \mathbf{0}'' \,\rangle$ is known. As a consequence, the recursive calculations will also require to treat all subsets of the current *set*, with dimensions ranging from 1 to $n-1$, so that $n \times n$ elements of $\mathbf{C}$ are effectively needed. When dealing with medium-to-large systems, the number of normal modes of the system can be really large while the classes remain comparatively small in most cases. From the previous observation, it means that in practice the program will need to look for a small quantity of elements in large vectors and matrices. From a computational perspective, this constant scan of the memory is largely inefficient and delays the calculations. To reduce the impact of this search on the speed of the procedure, new reduced matrices are created before starting the calculations by selecting the elements that will be needed in the original Sharp and Rosenstock matrices using the vector $\underline{\boldsymbol{v}}''_{\mathrm{id}}$. Figure 3.3 shows an example of making of a reduced matrix $\underline{\mathbf{C}}$ by extracting the required elements from the original matrix $\mathbf{C}$.

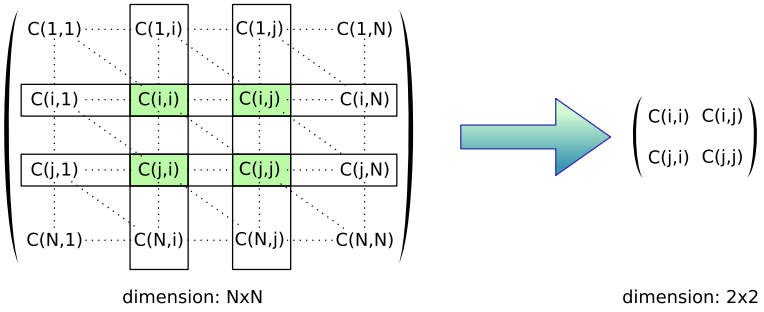


Figure 3.3: Construction of the matrix $\underline{\mathbf{C}}$ in the case of a *set* in class $\mathcal{C}_2$. The first excited mode is i and the second one is j, so that $\underline{v}''_{1_{\mathrm{id}}} = i$ and $\underline{v}''_{2_{\mathrm{id}}} = j$. In the same way, each element of the submatrix of $\mathbf{C}$ uses the same transformation. As an example, $\underline{\mathrm{C}}(1,2)$ satisfies the equivalence $\underline{\mathrm{C}}(1,2) = \mathrm{C}(\underline{v}''_{1_{\mathrm{id}}}, \underline{v}''_{2_{\mathrm{id}}})$.

Calculations of the transition dipole moment integrals assuming a linear variation of the electronic transition dipole moment with the normal coordinates have been partially discussed in section 2.8.1, but with an emphasis on the implications for the prescreening method. Here, we will focus on the possibilities of optimization of the computations and the use of submatrices to speed up the calculations. In the case of FCHT, the integrals are calculated with the formula:

$$\begin{aligned} \langle\, \mathbf{0}' \mid \boldsymbol{\mu}_{if} \mid \boldsymbol{v}'' \,\rangle = \boldsymbol{\mu}_{if}(\mathbf{Q}''_0) \langle\, \mathbf{0}' \mid \boldsymbol{v}'' \,\rangle + \sum_{k=1}^{N} \left(\frac{\partial \boldsymbol{\mu}_{if}}{\partial Q''_k} \right)_0 \sqrt{\frac{\hbar}{2\omega''_k}} \Big[\sqrt{v''_k}\, \langle\, \mathbf{0}' \mid \boldsymbol{v}'' - 1''_k \,\rangle \\ + \sqrt{v''_k + 1}\, \langle\, \mathbf{0}' \mid \boldsymbol{v}'' + 1''_k \,\rangle \Big] \end{aligned} \tag{3.12}$$

In the frame of a recursive calculation, equation 3.12 cannot be calculated immediately, as the N overlap integrals corresponding to the third term in its right-hand side are unknown at the time $\langle\, \mathbf{0}' \mid \boldsymbol{v}'' \,\rangle$ is evaluated. Using equation 2.38, it is possible to express these integrals with respect to overlap between lower quanta:

$$\begin{aligned}\langle\, \mathbf{0}' \mid \boldsymbol{\mu}_{if} \mid \boldsymbol{v}'' \,\rangle =& \boldsymbol{\mu}_{if}(\mathbf{Q}_0'') \langle\, \mathbf{0}' \mid \boldsymbol{v}'' \,\rangle + \sum_{k=1}^{N} \left(\frac{\partial \boldsymbol{\mu}_{if}}{\partial Q_k''}\right)_0 \sqrt{\frac{\hbar}{2\omega_k''}} \left[\sqrt{v_k''}\, \langle\, \mathbf{0}' \mid \boldsymbol{v}'' - 1_k'' \,\rangle \right. \\ &+ \sqrt{v_k''+1} \left\{ \frac{1}{\sqrt{2(v_k''+1)}} \Big[\mathrm{D}_k \langle\, \mathbf{0}' \mid \boldsymbol{v}'' \,\rangle \right. \\ &\left. \left. + \sum_{l=1}^{N} \sqrt{2v_l''}\, \mathrm{C}_{kl} \langle\, \mathbf{0}' \mid \boldsymbol{v}'' - 1_l'' \,\rangle \Big] \right\} \right]\end{aligned} \tag{3.13}$$

After transposing the general case to the treatment in *set* and reordering the terms in equation 3.13, we obtain:

$$\begin{aligned}\langle\, \mathbf{0}' \mid \boldsymbol{\mu}_{if} \mid \boldsymbol{v}'' \,\rangle =& \left[\boldsymbol{\mu}_{if}(\mathbf{Q}_0'') + \sum_{k=1}^{N} \left(\frac{\partial \boldsymbol{\mu}_{if}}{\partial Q_k''}\right)_0 \sqrt{\frac{\hbar}{2\omega_k''}} \frac{\mathrm{D}_k}{\sqrt{2}} \right] \langle\, \mathbf{0}' \mid \underline{\boldsymbol{v}}'' \,\rangle \\ &+ \sum_{k=1}^{n} \left[\left(\frac{\partial \boldsymbol{\mu}_{if}}{\partial \underline{Q}_k''}\right)_0 \sqrt{\frac{\hbar}{2\underline{\omega}_k''}} + \sum_{l=1}^{N} \left(\frac{\partial \boldsymbol{\mu}_{if}}{\partial Q_l''}\right)_0 \sqrt{\frac{\hbar}{2\omega_l''}} \mathrm{C}(l, \underline{v}_{k_{\mathrm{id}}}'') \right] \sqrt{\underline{v}_k''}\, \langle\, \mathbf{0}' \mid \underline{\boldsymbol{v}}'' - 1_k'' \,\rangle\end{aligned} \tag{3.14}$$

The factor before $\langle\, \mathbf{0}' \mid \underline{\boldsymbol{v}}'' \,\rangle$ is a constant that needs to be calculated only once. It is estimated along with $\langle\, \mathbf{0}' \mid \boldsymbol{\mu}_{if} \mid \mathbf{0}'' \,\rangle$ in the subroutine `FRANCK2` which handles the computations of the transition dipole moment integrals. It is then stored and re-used afterwards. It is also noteworthy that the derivatives of the electronic transition dipole moment are always coupled to the inverse square root of the reduced frequencies corresponding to the same mode. In actual calculation, it is more efficient to save and use directly the product of the derivatives and reduced frequencies:

$$\dot{\boldsymbol{\mu}}_k = \left(\frac{\partial \boldsymbol{\mu}_{if}}{\partial Q_k''}\right)_0 \sqrt{\frac{\hbar}{2\omega_k''}} \quad \text{and} \quad \underline{\dot{\boldsymbol{\mu}}}_k = \left(\frac{\partial \boldsymbol{\mu}_{if}}{\partial \underline{Q}_k''}\right)_0 \sqrt{\frac{\hbar}{2\underline{\omega}_k''}}$$

Finally, a wider submatrix of $\mathbf{C}$ than for FC is required here. Since the original submatrix $\underline{\mathbf{C}}$ is still needed for the recursion calculation of $\langle\, \mathbf{0}' \mid \underline{\boldsymbol{v}}'' \,\rangle$, increasing its size would mean a loss in efficiency. Another strategy adopted in FRANCK is to use a second submatrix, $\underline{\underline{\mathbf{C}}}$, to store the $n \times N$ elements needed in equation 3.14. However, instead of simply copying the elements, they are multiplied by the corresponding $\dot{\boldsymbol{\mu}}$, so that, for a mode i of the *set* corresponding to the mode j in the entire list of normal modes, we have the relation:

$$\underline{\underline{\mathrm{C}}}(k,i) = \dot{\boldsymbol{\mu}}_k \mathrm{C}(k, \underline{v}_{i_{\mathrm{id}}}'')$$

Using this submatrix, equation 3.14 can be written:

$$\langle\, \mathbf{0}' \mid \boldsymbol{\mu}_{if} \mid \boldsymbol{v}'' \,\rangle = \boldsymbol{\mu}_{\mathrm{cst}} \langle\, \mathbf{0}' \mid \underline{\boldsymbol{v}}'' \,\rangle + \sum_{k=1}^{n} \left[\underline{\dot{\boldsymbol{\mu}}}_k + \sum_{l=1}^{N} \underline{\underline{\mathrm{C}}}_{lk} \right] \sqrt{\underline{v}_k''}\, \langle\, \mathbf{0}' \mid \underline{\boldsymbol{v}}'' - 1_k'' \,\rangle \tag{3.15}$$

with

$$\boldsymbol{\mu}_{\text{cst}} = \boldsymbol{\mu}_{if}(\mathbf{Q}''_0) + \sum_{k=1}^{N} \left(\frac{\partial \boldsymbol{\mu}_{if}}{\partial Q''_k} \right)_0 \sqrt{\frac{\hbar}{2\omega''_k}} \frac{\mathrm{D}_k}{\sqrt{2}}$$

These schemes allow to reduce the number of operations to do when computing the intensity of each transition and so the computer costs of the FCHT/HT calculations. However, even with these improvements, they still are more cumbersome with respect to FC because of the $n \times N$ additional operations to perform for each computation of a transition intensity.

It is noteworthy that the creation of small matrices imply an increase of memory usage for the procedure. In case of FC calculations, this increase is really small as, for class $\mathcal{C}_{10}$ $\underline{\mathbf{C}}$ and $\underline{\mathbf{D}}$ will occupy $8 \times (10 \times 10 + 10) = 880$ B (B standing for bytes), which is very small on current computers. In the case of FCHT of HT calculation, the third submatrix $\underline{\underline{\mathbf{C}}}$ requires an additional memory proportional to the number of normal modes of the systems ($(8 \times 10 \times N)$ B for class $\mathcal{C}_{10}$). For information, this submatrix alone will require, in the case of class $\mathcal{C}_{10}$ about 1 kilobyte (1 KB = 2^{10} bytes) for a molecule of 6 atoms and 1 megabyte (1 MB = 2^{20} B) in the case of system with about 4371 atoms. The memory used by $\underline{\underline{\mathbf{C}}}$ can be reduced using the threshold ϵ_2^H presented in section 2.8 to remove the lowest contributions from the out-of-class modes in equation 3.15. On modern computer, however, the extra memory consumption induced by the usage of submatrices will not be noticeable, even with the entire submatrix $\underline{\underline{\mathbf{C}}}$ taken in account.

The calculations with the second derivatives of the electronic transition dipole moment are given in appendix A.

Until now, we described the generation of the *sets* with the use of nested loops. The problem of this approach is that the loops must be written explicitly. Hence, each class must be treated by an *ad hoc* procedure. A consequence of this approach is that it becomes very cumbersome to treat classes of higher order and it is practically impossible to devise a general-purpose method working for every class. For FRANCK to be as flexible as possible on its conditions of usage, generic loops were not used to treat classes $\mathcal{C}_3$ and above but a similar principle was used. It only requires beforehand the number of modes with non-zero maximum quantum numbers, $\underline{N}$, given by the prescreening and the order of the class, n. If n is greater than $\underline{N}$, the routine ends as not enough excited modes for the class were found by the evaluation method. Otherwise, the enumeration of the different combinations of final state is performed. Let us define Z_n the *set* to be calculated and $Z_{\underline{N}}$ the set of all modes chosen by the prescreening to be treated in a given class, so that $Z_n \subset Z_{\underline{N}}$. The first *set* is generated by taking the first n modes in $Z_{\underline{N}}$. Then, the first $(n-1)$ modes of Z_n are kept unchanged and the last one is replaced by the $(n+1)$-*th* mode in $Z_{\underline{N}}$. This sequence is reproduced until the last mode of $Z_{\underline{N}}$ has been used. At this moment, the second to last mode of Z_n, that is to say the $(n-1)$-*th* is replaced by the n-th mode in $Z_{\underline{N}}$ and the last mode of Z_n by the mode $(n+1)$ in $Z_{\underline{N}}$. As described before, the last mode of Z_n is regularly replaced until the last mode of $Z_{\underline{N}}$ has been treated. The entire procedure is repeated until Z_n corresponds to the last n modes in $Z_{\underline{N}}$, meaning that the loop is finished.

While simulating a arbitrary number of nested loops, the performance of such a procedure is lower than using real loops which can be optimized at the compiler level. However, most of the computational time is spent in calculation of the overlap integrals and the transition dipole moment integrals so that the increase of time was negligible in our tests with various systems, ranging from

phenoxyl radical to the coumarin C153 [82].

3.8 Managing the storage and the indexing of the overlap integrals

In the previous section, we focused on the possibility to efficiently compute the transition intensities considering a general approach independent from the prescreening method. We will discuss here about the technical implementation of the prescreening method presented in section 2.8 along with some considerations about the storage of the overlap integrals.

When the evaluation method was presented in section 2.8, it was explained then that this approach required three parameters, $\mathcal{C}_{1_{\max}}$, the maximum number of quanta to treat for each mode in $\mathcal{C}_1$, $\mathcal{C}_{2_{\max}}$, the maximum number of quanta for both simultaneously excited modes in $\mathcal{C}_2$, and $N_I^{\max}$ which represents an estimate of the maximum number of integrals to compute in each class. They can be set in our procedure with the keywords `MAXC1=`*`value`*, `MAXC2=`*`value`* and `MAXINT=`*`value`*, respectively. $\mathcal{C}_{1_{\max}}$ is set by default at 20 and $\mathcal{C}_{2_{\max}}$ at 13. `MAXINT` is expressed in millions of integrals, so that its default value is `MAXINT=100` which corresponds to $N_I^{\max} = 10^8$.

Classes $\mathcal{C}_1$ and $\mathcal{C}_2$ represent a limited number of possible final states to explore in a given *set* ($\mathcal{C}_{1_{\max}}$ for $\mathcal{C}_1$ and $(\mathcal{C}_{2_{\max}})^2$ for $\mathcal{C}_2$). As a consequence, it is possible to treat them with the Sharp and Rosenstock analytic formulae [43]. They are generated beforehand with a specialized program of symbolic algebra and directly inserted in the code of the procedure. A method similar to the one written by Weber and Hohlneicher [45] was used to obtain these formulae. Practically, it is impossible to cover all possible cases if $\mathcal{C}_{1_{\max}}$ and $\mathcal{C}_{2_{\max}}$ have been set really high by the user, such as it is required for SF_6^- described in section 4.3. However, it is possible to deal at least with the combinations treated with the default settings. Classes $\mathcal{C}_1$ and $\mathcal{C}_2$ have a specific function to compute the transition intensities. This one can handle equally a recursive or analytic calculation of the overlap integrals and the transition dipole moment integrals. By default, it attempts an analytic treatment and if the required formula is not found, falls back to the recursive calculations. From a computational perspective, on a modern computer (Opteron 2 GHz), the times to calculate the overlap integrals for the FC calculations, are roughly the same on a system such as the phenoxyl radical with a slight advantage for the analytic approach. The gain is more noticeable as the system becomes larger. The interest of the analytic formulae is twofold. Firstly, they avoid the risk of propagating the error in the calculation of an overlap integral such as it can happen with recursivity. This makes them very interesting to generate the pool of data that will be used later on for the prescreening. Secondly, in the case of FCHT and HT calculations, analytic formulae can directly provide the third term of equation 3.12 concerning the increase of one quantum. Computations are then slightly cheaper than a recursive treatment in this case. It is possible to deactivate the use of analytic formulae with the keyword `NOINTAN`.

A second specificity of the function used in classes $\mathcal{C}_1$ and $\mathcal{C}_2$ is that it is in charge of collecting the value that are required for the prescreening. While $F_{\mathcal{C}_1}$ and $F_{\mathcal{C}_2}$ are easily obtained, the values for $H_{\mathcal{C}_1}$, which only stores the HT contribution, are a bit more complex to evaluate in the case of recursion formulae because of our FC-HT mixing in $\boldsymbol{\mu}_{\text{cst}}$ in equation 3.15. As class $\mathcal{C}_1$ is very little demanding in terms of computational costs, it is easier to use $\boldsymbol{\mu}_{\text{cst}} - \boldsymbol{\mu}_{if}(\mathbf{Q}_0'')$ to generate $H_{\mathcal{C}_1}$ once the exact transition intensity has been calculated. Additionally, as discussed in reference [19], the value effectively stored for the evaluation of the HT contribution is $H_{\mathcal{C}_1}/I^{\text{tot}}$, so that we have to

compute:

$$H_{\mathcal{C}_1}(k, v_k'') = \frac{\left| \langle\, \mathbf{0}' \mid \boldsymbol{\mu}_{if} \mid \mathbf{0}'' + v_k'' \,\rangle - \boldsymbol{\mu}_{if}(\mathbf{Q}_0'')\langle\, \mathbf{0}' \mid \mathbf{0}'' + v_k'' \,\rangle \right|^2}{I^{\text{tot}}}$$

Currently, the implementation does not discard any terms outside the class so that the fourth array $H^{\text{oc}}_{\mathcal{C}_1}$ is not used. For higher classes starting from $\mathcal{C}_3$, the prescreening is called at the beginning. Depending on the *type of calculation*, FC or FCHT/HT, the prescreening follows the procedure described in section 2.8 or subsection 2.8.1, respectively. The three thresholds, ϵ_1, ϵ_2 and ϵ_1^H are all set equal to 10^{-9}. A single trial of the prescreening is computationally very cheap. Hence, the increment of the thresholds when the condition $N_I \leq N_I^{\max}$ is not met is set low: $\epsilon = 1.02 \times \epsilon$. Incidentally, the increment is the same for each threshold, which means that their values are always the same. A consequence of a slow evolution of the threshold is that a large number of trials might be needed to find a correct set $\boldsymbol{v}''_{\max}$ satisfying the condition on the number of integrals to estimate. This causes a difficulty for the output as it can become quickly really long. To avoid this problem, the current approach is to hide the prescreening. The reason for this choice, in addition to the large output generated, is that it will mix in the section dedicated to the assignement of the most likely transitions, making their reading more difficult. Because the prescreening itself does not provide any useful information for a common usage, this sequence can be safely hidden. However, a message can be written in the output during the prescreening:

```
Warning : insufficient number of quanta for mode i
```

This message means that the threshold was to low and that one of the conditions:

$$F_{\mathcal{C}_1}(i, v_i'') \geq \epsilon_1 \ , \quad H_{\mathcal{C}_1}(i, v_i'') \geq \epsilon_1^H (\text{for FCHT/HT calculation}) \ , \quad F_{\mathcal{C}_2}(i, j, v_i'') \geq \epsilon_2 \ (\forall j \neq i)$$

is met immediately for $v_i'' = \mathcal{C}_{1_{\max}}$ in the case of the first two conditions, or $v_i'' = \mathcal{C}_{2_{\max}}$ for the last one. Consequently, $\mathcal{C}_{1_{\max}}$ or $\mathcal{C}_{2_{\max}}$ should be increased to perform correctly the prescreening.

While the prescreening aims at improving the generation of the UV-vis spectra by correctly choosing the transitions to compute, the storage also plays an important role in the overall efficiency. As a matter of fact, in the recursion approach, a major part of the run time of the procedure is spent in the calculation of the transition dipole moment integrals on the one hand, and the retrieval of the overlap integrals for the recursion on the other hand. Storage is critical and must be kept as low as possible. An important reason discussed in 2.6 is that if not controlled carefully, it will quickly saturate the available memory. This problem is even more critical when tackling parallelization of the calculations. It was previously explained that classes starting from $\mathcal{C}_3$ could be treated contemporaneously. It is even possible to push further parallelization by handling simultaneously each *set*, that is to say each combination of excited modes of the final state in a class. When dealing with the *set* parallelization, the increase of memory is mainly due to the storage of the overlap integrals. In this case, a very high number of processors is needed to be really efficient, requiring a specific architecture. At this point, a *direct method*, without storage for the recursion can be even more adapted to this strategy. In our procedure, a simple partition in classes is used and the *sets* are then treated in serial. While some other arrays such as those containing the results

of the prescreening, need to be duplicated, the storage of the overlap integrals is the only one that can be easily limited. As hinted before, the lowest storage need was reached by managing the *sets* independently, so that the array dedicated to the overlap integrals is reused for each one. A consequence mentioned in the previous section is the necessity to recalculate all subsets contained in a given *set*. For a set Z_n, it will mean recomputing all sets of dimensions ranging from 1 to $n-1$. When considering a *set* from a higher class, it can represent a large number of recalculations, equal for a *set* Z_n to

$$\sum_{i=1}^{n-1} \frac{n!}{(n-i)!i!} .$$

For a *set* of class $\mathcal{C}_{10}$, it represents 1022 subsets previously treated to recalculate. These recalculations concern only the overlap integrals and not the transition dipole moment integrals and the line intensity. On a recent processor, this is not too cumbersome so the problem is manageable when the highest class is relatively low (inferior to 10).

Smaller arrays also mean faster retrieval of the elements inside. The latency to access to previously computed overlap integrals is critical in recursion formulae. As mentioned in section 2.6, the use here a minimal perfect hash function. This choice avoids any risk of collision, but the drawback is that the array must entirely fit in memory, otherwise storage is impossible. In practice, the maximum size occupied by the array among all possible *sets* is estimated once the prescreening has been performed and the list $\boldsymbol{v}''_{\max}$ is known. It is obtained for a class $\mathcal{C}_n$ as the product of the n highest maximum numbers of quanta incremented by one, to account for the subsets needed for the recursion. If the available memory is insufficient, the procedure will raise an error and ends. Otherwise, the calculations are performed normally. The size of the array is calculated for each *set* and corresponds to the number of overlap integrals to compute plus the overlap integral between the vibrational ground states, N_I^{set}:

$$N_I^{\text{set}} = \prod_{i=1}^{n} (\underline{v}''_{i_{\max}} + 1)$$

The calculations run on the indexes of the array from 1, which corresponds to $\langle\, \mathbf{0}' \mid \mathbf{0}'' \,\rangle$, to N_I^{set}. Two functions are then correlated to the storage, one to convert an array address to the corresponding vector of quantum numbers $\underline{\boldsymbol{v}}''$ and the inverse transformation given in equation 2.33. The latter is straightforward to implement while the first is slightly more complex and computationally more demanding because of test procedures. As hinted by the hash function to convert the quantum numbers into an array index, the counting of the mode is as follows. The quantum numbers of the first mode is incremented until it reaches $\underline{v}''_{1_{\max}}$. Then, the number of quanta of the second mode is incremented by one ($\underline{v}''_2 = 1$) and the quantum number of the first mode is reset to zero. It is then incremented again until it equals the limit $\underline{v}''_{1_{\max}}$. The sequence is run until the quantum numbers of the n modes reach their maximum. Theoretically, for any index i, one just need to performs the procedures until it has reached $i-1$ operation, the first index being the starting point of the enumeration. In practice, it is rather time consuming and not efficient. It is possible to reduce the time of the process by half (in the case of FC calculations with phenoxyl radical) by using some observations about the counting, for a given index i:

- if $i \leq \underline{v}''_{1_{\max}} + 1$, then the final state is $|\, \underline{\mathbf{0}}'' + (i-1)''_1 \,\rangle$.

- if $i = \prod_{\underline{i}=1}^{\underline{n}}(\underline{v}''_{\underline{i}_{\max}} + 1)$, with $\underline{n} \leq n$, then the final state is $|\, \underline{\mathbf{0}}'' + \sum_{\underline{i}=1}^{\underline{n}} \underline{v}''_{\underline{i}_{\max}} \,\rangle$[g].
- if $i = \prod_{\underline{i}=1}^{\underline{n}}(\underline{v}''_{\underline{i}_{\max}} + 1) + 1$, with $\underline{n} < n$, then the final state is $|\, \underline{\mathbf{0}}'' + 1''_{\underline{n}+1} \,\rangle$.

If none of this condition is met for a given index, then the quantum number of the highest excited mode $\underline{n}$ is obtained by the relation:

$$\underline{v}''_{\underline{n}} = \left\lfloor \frac{i}{\prod_{\underline{i}=1}^{\underline{n}-1}(\underline{v}''_{\underline{i}_{\max}} + 1)} \right\rfloor$$

The index i is then modified:

$$i = i - \underline{v}''_{\underline{n}} \times \prod_{\underline{i}=1}^{\underline{n}-1}(\underline{v}''_{\underline{i}_{\max}} + 1)$$

The new index is confronted again to the conditions stated above, this time on the subset $\underline{n}-1$. If none is met, $\underline{v}''_{\underline{n}-1}$ is evaluated in the same way as previously. The procedure is repeated until the quantum number of the first mode has been found.

3.9 Conclusion and discussion

In this chapter, we focused on the technical implementation of the methods presented in chapter 1 and discussed more in details in chapter 2. While the development of a consistent and versatile method to compute the overlap integrals and the transition intensities is very important, the way it is encoded cannot be overlooked at the risk otherwise to lose greatly in efficiency of the resulting program. Consequently, a great care has been taken to optimize the procedure implemented inside GAUSSIAN and to simplify the implementation of new features or algorithms.

As an example, in the current version, the temperature is not fully functional so it has not been treated here. The discussions remain true in the case of vibrational excited states in the initial state as they can be simply duplicated to the treatment of the initial state. The vibronic initial states to consider can be chosen with respect to their Boltzmann population and a similar system of *sets* can be applied to them. One of the difficulty lies on the fact that the memory usage is noticeably increased for the storage of the overlap integrals, which can be problematic for large systems. Another issue is that the optimization of the computations is more complex as two recursion formulae are required. Apart from the matter of efficient calculations, the structure could simply integrate the management of the temperature by introducing in the second subroutine `FRANCK2` shown in figure 3.1 a loop to perform the calculations of the transition dipole moment integrals for each *set* of the initial state.

Another current limitation is that the electronic transition dipole moment and its derivatives are expected to be given with respect to the final state. Indeed, considering the Taylor series of $\boldsymbol{\mu}_{if}$ about the equilibrium geometry of the initial state can be more complex, depending on the *type of calculation* required. Applying strictly the Franck-Condon principle, the electronic transition

[g] $\underline{v}''_{\underline{i}_{\max}}$ is the quantum number of the $\underline{i}$-*th* mode of the *set*. To avoid too many subscripts, it was not explicitly written in the ket.

dipole moment is constant:

$$\langle\, \boldsymbol{v}' \mid \boldsymbol{\mu}_{if} \mid \boldsymbol{v}'' \,\rangle = \boldsymbol{\mu}_{if}(\mathbf{Q}'_0)\langle\, \boldsymbol{v}' \mid \boldsymbol{v}'' \,\rangle \tag{3.16}$$

It is noteworthy that, if the relation $\boldsymbol{\mu}_{if}(\mathbf{Q}'_0) = \boldsymbol{\mu}_{if}(\mathbf{Q}''_0)$ is not satisfied, the absolute intensities of the spectral bands are different depending on whether the transition dipole moment integrals are computed with respect to the initial or final state, but the ratio of height between two peaks is the same. Such a problem would not be visible in the result given by FRANCK since the electronic transition dipole moment is currently disregarded in this case. However, if one wishes to compare different spectra, this difference should be taken into account.

When extending the Taylor series of $\boldsymbol{\mu}_{if}$ given in equation 2.42 up to the first order, calculations become more difficult. For FCHT calculation, the transition dipole moment integral is:

$$\langle\, \boldsymbol{v}' \mid \boldsymbol{\mu}_{if} \mid \boldsymbol{v}'' \,\rangle = \boldsymbol{\mu}_{if}(\mathbf{Q}'_0)\langle\, \boldsymbol{v}' \mid \boldsymbol{v}'' \,\rangle + \langle\, \boldsymbol{v}' \mid \sum_{k=1}^{N} \left(\frac{\partial \boldsymbol{\mu}_{if}}{\partial Q'_k}\right)_0 Q'_k \mid \boldsymbol{v}'' \,\rangle \tag{3.17}$$

The problem here is similar to the one we encountered when calculating the analytic limit $I^{\text{tot}}_{\mathbf{Q}''}$ in section 2.9. Using the Duschinsky transformating given in equation 1.64, equation 3.17 can be written:

$$\langle\, \boldsymbol{v}' \mid \boldsymbol{\mu}_{if} \mid \boldsymbol{v}'' \,\rangle = \boldsymbol{\mu}_{if}(\mathbf{Q}'_0)\langle\, \boldsymbol{v}' \mid \boldsymbol{v}'' \,\rangle + \langle\, \boldsymbol{v}' \mid \sum_{k=1}^{N} \left(\frac{\partial \boldsymbol{\mu}_{if}}{\partial Q'_k}\right)_0 \left(\sum_{l=1}^{N} \mathrm{J}_{kl} Q''_l + \mathrm{K}_l\right) \mid \boldsymbol{v}'' \,\rangle \tag{3.18}$$

$$\begin{aligned} &= \left[\boldsymbol{\mu}_{if}(\mathbf{Q}'_0) + \sum_{k=1}^{N} \left(\frac{\partial \boldsymbol{\mu}_{if}}{\partial Q'_k} \sum_{l=1}^{N} \mathrm{K}_l\right)_0\right] \langle\, \boldsymbol{v}' \mid \boldsymbol{v}'' \,\rangle \\ &\quad + \sum_{k=1}^{N} \left(\frac{\partial \boldsymbol{\mu}_{if}}{\partial Q'_k}\right)_0 \sum_{l=1}^{N} \mathrm{J}_{kl} \langle\, \boldsymbol{v}' \mid Q''_l \mid \boldsymbol{v}'' \,\rangle \end{aligned} \tag{3.19}$$

Finally, we replace the normal coordinates of the final state Q''_l by the annihilation and creation operators using the relation given in equation 1.53:

$$\begin{aligned} \langle\, \boldsymbol{v}' \mid \boldsymbol{\mu}_{if} \mid \boldsymbol{v}'' \,\rangle = &\left[\boldsymbol{\mu}_{if}(\mathbf{Q}'_0) + \sum_{k=1}^{N} \left(\frac{\partial \boldsymbol{\mu}_{if}}{\partial Q'_k} \sum_{l=1}^{N} \mathrm{K}_l\right)_0\right] \langle\, \boldsymbol{v}' \mid \boldsymbol{v}'' \,\rangle \\ &+ \sum_{k=1}^{N} \left(\frac{\partial \boldsymbol{\mu}_{if}}{\partial Q'_k}\right)_0 \sum_{l=1}^{N} \mathrm{J}_{kl} \left(\frac{\hbar}{2\omega_l}\right)^{1/2} \\ &\times \left[\sqrt{v''_l}\, \langle\, \boldsymbol{v}' \mid \boldsymbol{v}'' - 1''_l \,\rangle + \sqrt{v''_l + 1}\, \langle\, \boldsymbol{v}' \mid \boldsymbol{v}'' + 1''_l \,\rangle\right] \end{aligned} \tag{3.20}$$

We find now a familiar albeit slightly more complex formula with respect to equation 2.37. Contrary to what seems to show equation 3.20, computational costs are on par with the calculations of the transition dipole moment integrals with respect to the final state. To justify this statement,

it is enough to reorder the sums of the second term in the right-hand side of the equation:

$$\langle\, \boldsymbol{v}' \mid \boldsymbol{\mu}_{if} \mid \boldsymbol{v}'' \,\rangle = \boldsymbol{\mu}_{\text{cst}} \langle\, \boldsymbol{v}' \mid \boldsymbol{v}'' \,\rangle + \sum_{k=1}^{N} \dot{\boldsymbol{\mu}}_k \left[\sqrt{v_k''}\, \langle\, \boldsymbol{v}' \mid \boldsymbol{v}'' - 1_k'' \,\rangle + \sqrt{v_k'' + 1}\, \langle\, \boldsymbol{v}' \mid \boldsymbol{v}'' + 1_k'' \,\rangle \right] \quad (3.21)$$

with

$$\boldsymbol{\mu}_{\text{cst}} = \boldsymbol{\mu}_{if}(\mathbf{Q}_0') + \sum_{k=1}^{N} \left(\frac{\partial \boldsymbol{\mu}_{if}}{\partial Q_k'} \sum_{l=1}^{N} \mathrm{K}_l \right)_0$$

$$\dot{\boldsymbol{\mu}}_k = \left(\frac{\hbar}{2\omega_k} \right)^{1/2} \sum_{l=1}^{N} \mathrm{J}_{lk} \left(\frac{\partial \boldsymbol{\mu}_{if}}{\partial Q_l'} \right)_0$$

$\boldsymbol{\mu}_{\text{cst}}$ and the N elements $\dot{\boldsymbol{\mu}}_k$ can be computed beforehand once and stored to be used directly in calculations. Using a factorization method as in equation 3.15 is perfectly feasible. The drawback of this solution is to rely on the Duschinsky transformation to transform the normal coordinates of the initial states. This is not really a major problem since all our calculations of the transition dipole moment integrals suppose that this approximation is accurate. However, for medium-to-large systems, this approximation is often very good.

When considering spectra generated with an approximation of the electronic transition dipole moment with respect to the initial or final state, it is noteworthy that differences can be seen depending on the quality of the description of $\boldsymbol{\mu}_{if}$ by the chosen Taylor series. We have already mentioned that if the FC approximation is insufficient to describe correctly $\boldsymbol{\mu}_{if}$ ($\boldsymbol{\mu}_{if}(\mathbf{Q}_0') \neq \boldsymbol{\mu}_{if}(\mathbf{Q}_0'')$) then the spectral bands will have different heights depending on the state of reference. More generally, it is straightforward to see that if the Taylor expansion carried out up to a given order is not accurate enough, then the spectra generated using equations 1.29 or 2.42 will show discrepancies.Additionally, even if this approximation is precise, one should be careful when analyzing the contributions corresponding to each order of the Taylor series in the generated spectra. To illustrate this, let us assume that the Herzberg-Teller approximation gives an exact description of the electronic transition dipole moment, so that we have the following relation:

$$\boldsymbol{\mu}_{if} = \boldsymbol{\mu}_{if}(\mathbf{Q}_0') + \sum_{k=1}^{N} \left(\frac{\partial \boldsymbol{\mu}_{if}}{\partial Q_k'} \right)_0 Q_k' = \boldsymbol{\mu}_{if}(\mathbf{Q}_0'') + \sum_{k=1}^{N} \left(\frac{\partial \boldsymbol{\mu}_{if}}{\partial Q_k''} \right)_0 Q_k'' \quad (3.22)$$

Applying this equality to the calculation of the transition dipole moment integrals, one obtains:

$$\begin{aligned} &\boldsymbol{\mu}_{if}(\mathbf{Q}_0') \langle\, \boldsymbol{v}' \mid \boldsymbol{v}'' \,\rangle + \sum_{k=1}^{N} \left(\frac{\partial \boldsymbol{\mu}_{if}}{\partial Q_k'} \right)_0 \langle\, \boldsymbol{v}' \mid Q_k' \mid \boldsymbol{v}'' \,\rangle \\ &\quad = \boldsymbol{\mu}_{if}(\mathbf{Q}_0'') \langle\, \boldsymbol{v}' \mid \boldsymbol{v}'' \,\rangle + \sum_{k=1}^{N} \left(\frac{\partial \boldsymbol{\mu}_{if}}{\partial Q_k''} \right)_0 \langle\, \boldsymbol{v}' \mid Q_k'' \mid \boldsymbol{v}'' \,\rangle \end{aligned} \quad (3.23)$$

Since we assumed that the electronic transition dipole moment was correctly described by an expansion in a Taylor series up to the first order, it is expected that the FC principle was unsufficient, which means that $\boldsymbol{\mu}_{if}(\mathbf{Q}_0') \neq \boldsymbol{\mu}_{if}(\mathbf{Q}_0'')$. The FC contribution to the transition intensities correspond to the squared norm of the first terms in both left-hand side and right-hand side of equation 3.24, while the HT contribution is given by the second terms. Since $\boldsymbol{\mu}_{if}(\mathbf{Q}_0') \neq \boldsymbol{\mu}_{if}(\mathbf{Q}_0'')$, it is

straightforward to see that

$$\left|\boldsymbol{\mu}_{if}(\mathbf{Q}_0'')\right|^2\left|\langle\, \boldsymbol{v}' \mid \boldsymbol{v}'' \,\rangle\right|^2 \neq \left|\boldsymbol{\mu}_{if}(\mathbf{Q}_0')\right|^2\left|\langle\, \boldsymbol{v}' \mid \boldsymbol{v}'' \,\rangle\right|^2$$

which means that the FC contributions will not be the same if the spectrum is generated using equation 1.29 or 2.42 as the description of μ_{if}. For the HT, contribution, let us rewrite equation 3.24 in the following form:

$$\begin{aligned}\sum_{k=1}^{N}\left(\frac{\partial\boldsymbol{\mu}_{if}}{\partial Q_k'}\right)_0 \langle\, \boldsymbol{v}' \mid Q_k' \mid \boldsymbol{v}'' \,\rangle = \left[\boldsymbol{\mu}_{if}(\mathbf{Q}_0'') - \boldsymbol{\mu}_{if}(\mathbf{Q}_0')\right]\langle\, \boldsymbol{v}' \mid \boldsymbol{v}'' \,\rangle \\ + \sum_{k=1}^{N}\left(\frac{\partial\boldsymbol{\mu}_{if}}{\partial Q_k''}\right)_0 \langle\, \boldsymbol{v}' \mid Q_k'' \mid \boldsymbol{v}'' \,\rangle\end{aligned} \tag{3.24}$$

Since the first term in the right-hand side of equation 3.24 is not null, the HT contributions to the transition intensities are different depending on whether the Taylor expansion of μ_{if} has been carried out with respect to the equilibrium geometry of the initial of final state. Consequently, while the spectrum of the FC contribution will only differ by the absolute intensities of the spectral bands, in the case of the HT contribution, even the shape of the spectrum will be different.

As a matter of fact, and as hinted in section 2.9, it should be better to perform the Taylor expansion of the electronic transition dipole moment about the equilibrium geometry of the initial state. Such a choice is also more in line with the perturbation theory.

In addition to considerations related to the code optimization, the possibilities of customization given to the user were presented. Some of these options were necessary to carry out the calculations, such as those spectrum-related, or to adapt to particular cases, such as the parameters of prescreening. Others have been added from experience gained with the application of the procedure to study different systems, such as the mode-specific scaling scheme. A list of the keywords available in the procedure is given in appendix E. We will see some cases studied during this work in the next chapter.

Chapter 4

Examples of application of the procedure

4.1 Introduction

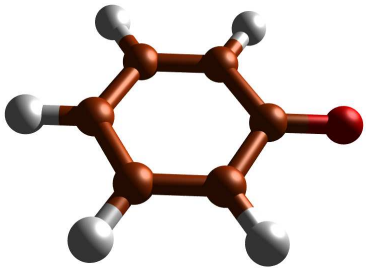
Once the procedure inside the computational package has been mature enough, it could be used on real cases. Some of them have been chosen to present the possibilies offered by the embedding inside a general-purpose quantum chemistry package such as GAUSSIAN, and others are practical studies to analyze of experimental results [19, 20].

In this chapter, we will present several of these systems. The first one, the phenoxyl radical, has been extensively used throughout this work to test the procedure and check new developments. It was also used as a benchmark when confronting different implementations of the same theoretical method. We will present with this model system a generic usage of the procedure and the results obtained.

It will be followed by several short studies designed to illustrate the influence of some parameters on the generated final spectrum. For sulfur hexafluoride, the defaul parameter for $\mathcal{C}_{1_{\max}}$ is too low, leading to a truncated spectrum. Chlorophyll *c2* is a large system with many normal modes. It will be used to discuss the spectrum convergence, with respect to the analytic limit, but also with respect to the spectrum shape.

As mentioned before, the integration inside a quantum chemistry package gives a straightforward access to different kinds of computations. As examples, we will show the cases of acrolein in a solvent environment and adenine adsorbed on a cluster of silicon. The latter is also an interesting application of our procedure to a large system.

To conclude, we will present two practical studies. In the case of anisole, the vibrationally-resolved absorption spectrum obtained with the Franck-Condon calculations was used to propose a new assignment to the spectral bands of a REMPI spectrum. Finally, as an introduction to chlorophyll *c2*, we will present a study of the absorption and emission spectra of the porphyrin [19].



4.2 Phenoxyl

4.2.1 Introduction

Phenoxyl radical plays an important role in combustion chemistry and biology. As an example, it is a key intermediate, used alone or in a metal complex, in biocatalysis, protein redox reactions, electron-transfer reactions and biosynthesis (see reference [104] and references therein). It is also useful as a prototype to discuss the chemical and electronic properties of larger aromatic systems containing oxygen.

Recently, Dierksen and Grimme [93] used this system as a test case for their algorithm of block-diagonalization [70]. This molecule represents an interesting challenge from a computational and as well as spectroscopic perspective. For the latter, the difficulty comes from its high reactivity toward both second order combination[a] and electron transfer processes. As a consequence, it can only be observed for a short time. It is possible to sufficiently stabilize the phenoxyl radical by choosing an adequate precursor and working at a very low temperature. The analysis of the electronic absorption spectrum still remains complex and has been subject to different interpretations as described in reference [105].

We will focus on the weakly-allowed $\pi - \pi^*$ transition between the electronic ground state which has a $^2B_1$ symmetry and final state with a $^2A_2$ symmetry. Being weakly-allowed, it represents an interesting case to study the effects of the Franck-Condon and Herzberg-Teller approximations on the final spectrum. However, comparison with an experimental absorption spectrum is rendered difficult by the slight discrepancies observed for the same transition depending on the precursor of the phenoxyl. Despite this fact, the molecule was chosen as a reference system to test our procedure during its development and to test new functionalities for several reasons. Thanks to its aromatic ring, the structure is quite rigid so that the geometry remains mostly unchanged during the electronic transition. The system is also a good compromise in terms of number of vibrations, being sufficiently small to allow quick tests while large enough to provide meaningful results when checking the influence of parameters or the correct behavior of the entire procedure.

In this section, we will mostly discuss the usage of FRANCK inside GAUSSIAN, to the expense of a detailed analysis of the results. This example can be regarded as an application of the theoretical presentation given in chapter 3 and we will regularly refer to details explained in the latter.

4.2.2 Computational details

From a theoretical perspective, the phenoxyl radical also represents a difficult case for electronic structure methods because it requires a sophisticated treatment of the electron correlation. For our

[a] A reaction between phenoxyl radicals when their concentration exceeds 10^{-4} M, forming new products such as 2,2'- 2,4' and 2,6'-dihydroxybiphenyls.

need, the unrestricted TD-DFT method provided very satisfying results and was able to describe correctly the excited state [12] while DFT was used for the ground state. The B3LYP exchange-correlation functional and the TZVP basis set were employed.

As described in section 3.3, a standard usage of Franck requires four basic steps, two for the geometry optimization and two for the evaluation of the frequencies. The generation of the requested spectrum is performed in the last GAUSSIAN job, once the necessary data have been calculated. As mentioned before, it is advisable to add a fifth step corresponding to the backup of the checkpoint file to avoid the overwriting of the data obtained at the end of step 3. The complete sequence is reproduced here to facilitate further references:

1. Optimization of the geometry of state 1
2. Optimization of the geometry of state 2
3. Frequencies calculation for state 1
4. Backup of the checkpoint file obtained in step 3
5. Frequencies calculation for state 2 + generation of the spectrum

As a remark, state 1 and state 2 can be indifferently the initial and final states. By default, state 1 is assumed to be the initial one and state 2 the final one. Steps 1 and 3 can be run together in a single GAUSSIAN job but not steps 2 and 4 since geometry optimization causes an overwriting of the checkpoint file during runtime.

Step 4 allows to generate several spectra with the same input data but different parameters, by performing each time only the last step[b]. In such a case, it is advisable to handle the state with the more demanding calculations first. Once the frequencies have been computed and all necessary data stored in the checkpoint file (end of step 3), the latter can be recalled for several tests without redoing steps 1 to 4.

In section 3.3, we also mentioned the possibility to use GAUSSIAN output files as data sources for the initial and/or final states. It should be noted that GAUSSIAN still expects a job to perform even if one does not wish to use information from the calculation in progress to generate the UV-vis spectrum. Moreover, FRANCK always extracts some data from GAUSSIAN, such as the number of atoms and normal modes and the atomic masses. Consequently, it is necessary to use the same molecule for the job and in the input files for the generation of the spectrum. However, if only the spectrum is desired, one can choose a very "light" job such as a single point energy calculation with a molecular mechanics method (UFF for example).

In this study, we will only consider a standard usage of the routine (GAUSSIAN calculation and checkpoint file as data sources for the spectrum generation) and assume that the geometry optimization has already been performed for both states, meaning that steps 1 and 2 have been done. The input files for step 3 and 5 are given in section F.1. In the configuration used in this example, state 1 was the final state and state 2 the initial state.

[b] To be thorough, we should mention that it is even possible to start almost directly from the calculations of the Franck-Condon spectra using the *read-write files* and the subprogram l716.exe. However, such a procedure is rather complex and cumbersome, and requires a good knowledge of GAUSSIAN. We will restrict our explanations to a simple usage.

4.2.3 Discussion

The weakly-allowed transition $^2B_1 \rightarrow {}^2A_2$ is the lowest optically observable electronic one in energy. Photoelectron spectroscopy has shown that an electronic transition also existed in the near-IR region [106] and theoretical calculations have been able to predict it [92]. However, optical and polarization techniques have been unable to characterize it [105]. As mentioned in the introduction, we restricted our study to a single electronic transition. The computed TD-B3LYP,B3LYP/TZVP energy of transition equals 17000 cm^{-1}, slightly superior to the expected 16000 cm^{-1} [105]. The vibrational frequencies are listed in table 4.2. In the second and third columns are reported the calculated frequencies with the harmonic and anharmonic approximation and in the fourth column the experimental frequencies. In columns 6 and 7, the normal modes of the final states are listed by increasing energy. The main terms of their projections on the normal coordinates of the final state are reported in column 6. The coefficients correspond to the squared elements of the Duschinsky matrix, $J(v_i', v_j'')^2$.

We will now detail the generation of the absorption spectrum with the Franck-Condon approximation. As mentioned in section 3.2, the keyword `FC` is used in the *dinautil section*. In our example, the data for the final state are extracted from the checkpoint file, which is not the expected setting. As a consequence, we must give explicitly the correct source with the parameter `CHK2`. In theory, it is unnecesssary in our case to specify the second source, since the procedure will automatically choose the current GAUSSIAN job. However, it is advisable to give both input when performing a "non-standard" calculation to avoid some unexpected behavior. Finally, we require the printing of the Duschinsky matrix **J** and the shift vector **K**. The former has been used for our assignment of the final state vibrations (see table 4.2) and is also very useful to control the mode mixing. A large coupling can cause the prescreening method to lose partly its efficiency. The shift vector is interesting to control if the displacement of the nuclei during the transition is not too important. The *dinautil section* that we will use is:

```
NoReord FC CALC1 CHK2 PRTMAT=12
```

It should be noted that we voluntarily rotated the structure of the initial state to be at an angle with the one of the final state. This was done to test the superposition procedure in FRANCK. The reorientation is automatic and carried out as soon as all input data have been extracted from the respective sources (the reoriented structures are given in section F.2). When finished, the nuclear displacements induced by the vibrations are recalculated as well as the electronic transition dipole moment and its derivatives. FRANCK will then print the requested matrices, here **J** and **K** and start the calculation of the transition intensities to generate the spectrum.

With the default settings, the procedure takes 136 seconds on an AMD Opteron™ 2.4 GHz. The spectrum progression (SP), reported in the output as a conclusion of the treatment of each class, represents the recovery of the total spectrum intensity given by the analytic limit described in section 2.9. The total spectrum convergence, as well as the specific progression for each class are shown in table 4.1.

[c] $\Delta_{SP} = SP_n - SP_{n-1}$

We obtain a very good convergence of the spectrum intensity. This is an expected result because of the rigidity and the C_{2v} symmetry of the molecule. The Duschinsky matrix in figure 4.1 shows a very limited mode-mixing where our prescreening method is particularly efficient.

For a convergence of 99%, it is sufficient to go up to class $\mathcal{C}_5$. Setting `MAXBANDS=5`, the computational time is reduced to 62 seconds. This difference is even stronger when considering FCHT calculations (same *dinautil section* except for the `FCHT` keyword instead of `FC`). In this case, default calculations are done in 16 min 26 s while it takes "only" 6 min 37 s up to class $\mathcal{C}_5$. This trend is obviously more pronunced when dealing with large systems and avoiding large classes when unnecessary can greatly reduce the computational costs.

Class	Convergence	$\Delta_{SP}{}^{c}$
0	3.57	3.57
1	28.43	24.86
2	63.12	34.69
3	87.08	23.96
4	96.87	9.79
5	99.43	2.56
6	99.83	0.40
7	99.87	0.04

Table 4.1: Spectrum convergence (given in %) for the absorption spectrum of the phenoxyl radical with the FC approximation

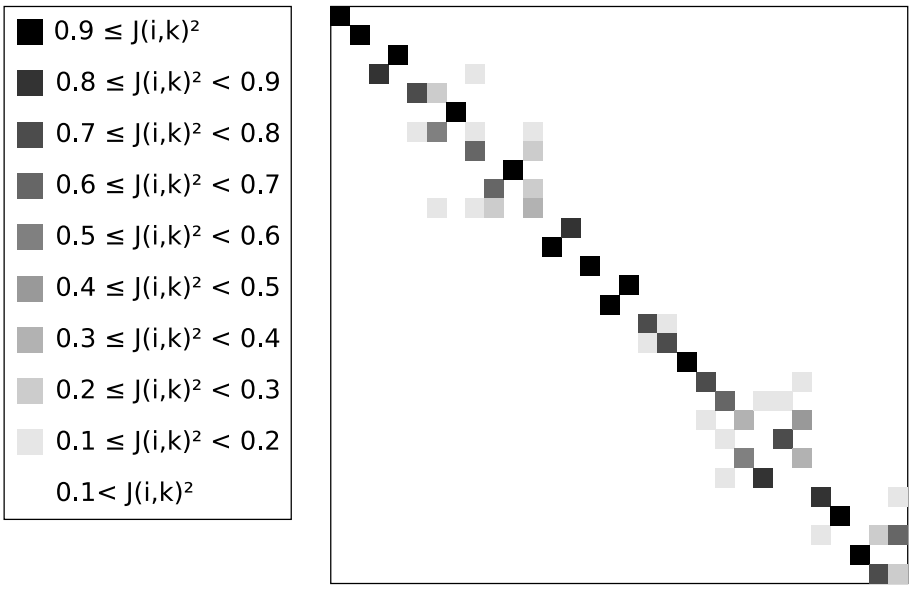


Figure 4.1: Squared elements of the Duschinsky matrix $\mathrm{J}(v_i', v_j'')^2$.

To compare with the experimental spectrum taken from reference [105], the bands of the spectrum generated with the default settings (`SPECHWHM=135.`), are too large. Hence, we reduce the half-width at half-maximum for the convolution to 100 cm^{-1}. Moreover, it is possible to limit the upper bound of the spectrum to 4500 cm^{-1} with respect to the energy of transition between the vibrational ground states. The *dinautil section* becomes:

```
NoReord FC CALC1 CHK2 PRTMAT=12 SPECHWWM=100. SPECMAX=4500.
```

As an illustration of our previous discussion in section 3.4 about the influence of the spectrum parameters on the calculation times, the procedure takes 109 s for FC and 15 min 56 s for FCHT to deal with the computations of the transition intensities up to $\mathcal{C}_7$. Two remarks can be done

on this improvement. Firstly, since the changes are relatively limited, the gain in computational time is relatively low. Secondly, the latter is almost equivalent in both cases (FC or FCHT). Such a behaviour is expected since the number of transitions actually computed is relatively similar (limited by $N_I^{\max}$).

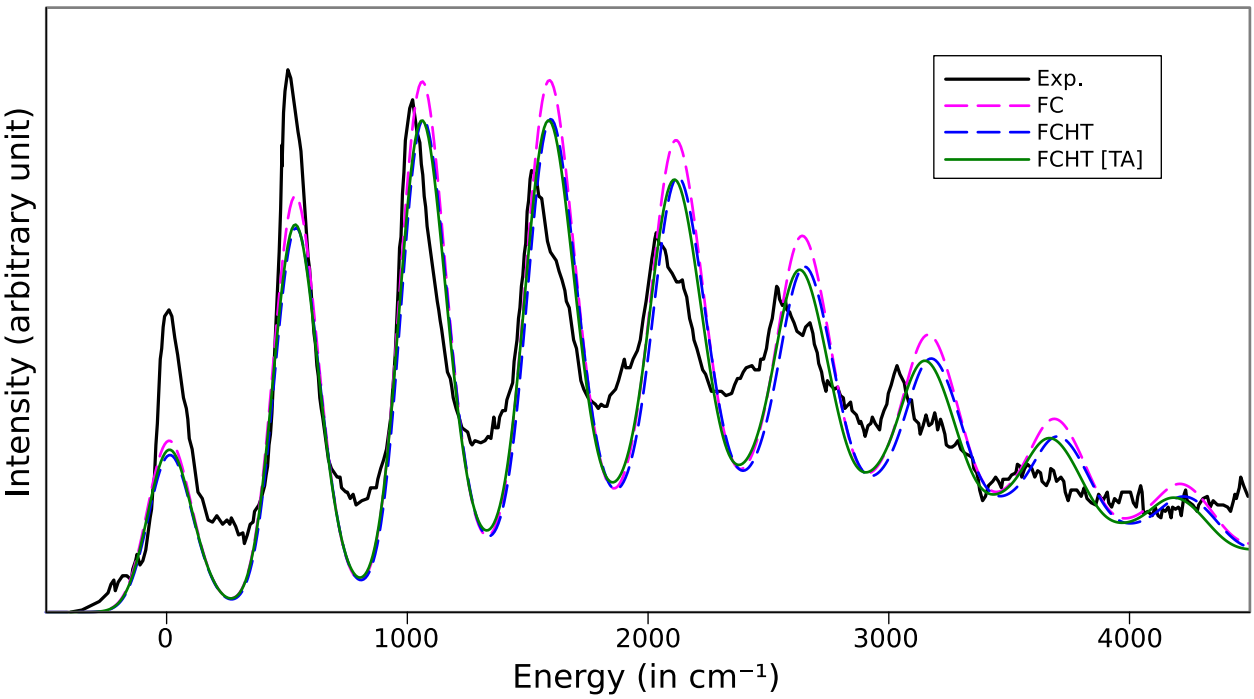


Figure 4.2: Experimental and theoretical spectra of the $^2B_1 \rightarrow {}^2A_2$ transition of the phenoxyl radical. The experimental spectrum from reference [105] is represented in solid black line. Magenta and blue dashed lines represent the FC and FCHT spectra obtained from harmonic calculations, respectively. The solid green line shows the FCHT spectrum obtained with the mode-specific scaling of anharmonicity described in section 3.3.1. Half-width at half-maximum is 100 cm^{-1} for all theoretical spectra.

The FC and FCHT spectra obtained with these settings are respectively shown with a dashed blue and magenta line in figure 4.2. To allow comparison with the experimental spectrum shown in the same figure with a solid black line, the energy of the 0-0 transition has been set to the origin, so that the energies of the spectral bands are given in relative values. Despite some differences in the band intensities, the main peaks are correctly reproduced and we can observe an improvement of the accuracy by taking into account the Herzberg-Teller approximation. Nonetheless, we find a small blue shift of the bands of higher energies. To try to correct this, we use the anharmonic scaling scheme described in section 3.3.1 using anharmonic frequencies calculated with the N07 basis set [107] and reported in the third column of table 4.2. The *dinautil section* for the FCHT spectra is now:

```
 NoReord FCHT CALC1 CHK2 PRTMAT=12 SPECHWWM=100. SPECMAX=4500. SclVec

 184.073 371.406 443.384 474.178 524.348 589.958 641.322 785.018 788.014
 795.326 905.604 973.180 963.945 976.118 994.008 1075.446 1152.041 1152.759
 1256.681 1323.310 1395.345 1415.318 1462.820 1515.760 1562.620 3046.396
 3058.730 3071.912 3082.009 3082.952
```

The generated spectrum with the anharmonicity scaling (FCHT [TA]) is reported with a solid green line in figure 4.2. We can see a small shift of the bands at higher energies.

As described in section 3.4, the assignment for the most intense bands (chosen with the keyword `PRTINT`) is written in the output. As an example, the following information are given for some of the most probable transitions in the case of FCHT [TA]:

```
Initial State: <0|
Final State: |0>
     DeltaE =     0.0000 | TDMI**2 = 0.4196E-02, Intensity = 0.4005E-02
   ....................................
Initial State: <0|
Final State: |5^1>
     DeltaE =   517.4194 | TDMI**2 = 0.7899E-02, Intensity = 0.7754E-02
   ....................................
Initial State: <0|
Final State: |5^2>
     DeltaE =  1034.8389 | TDMI**2 = 0.7128E-02, Intensity = 0.7216E-02
   ....................................
Initial State: <0|
Final State: |5^3>
     DeltaE =  1552.2583 | TDMI**2 = 0.4100E-02, Intensity = 0.4272E-02
   ....................................
[...]
Initial State: <0|
Final State: |6^1;5^1>
     DeltaE =  1056.2777 | TDMI**2 = 0.3651E-02, Intensity = 0.3703E-02
   ....................................
[...]
Initial State: <0|
Final State: |6^1;5^2>
     DeltaE =  1573.6971 | TDMI**2 = 0.3654E-02, Intensity = 0.3804E-02
   ....................................
```

Using the assignment of this spectrum, we generated the stick spectrum with an external script. The relative intensities of the main peaks are very similar to the experimental spectrum.

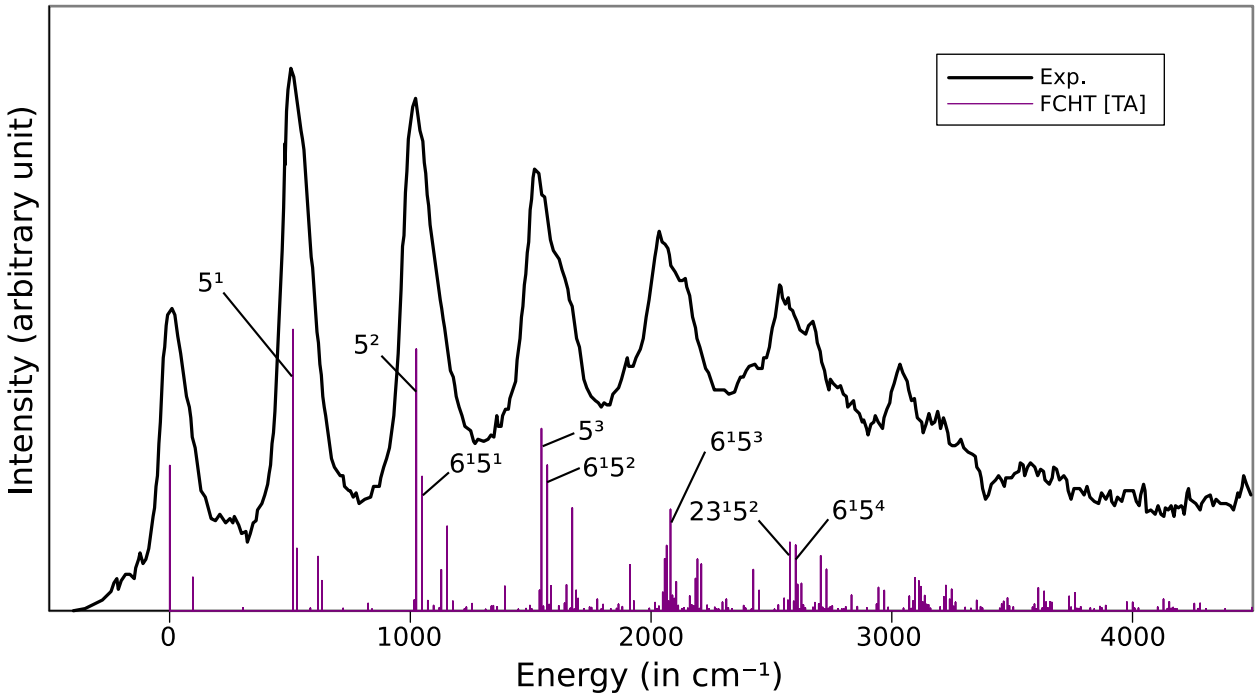
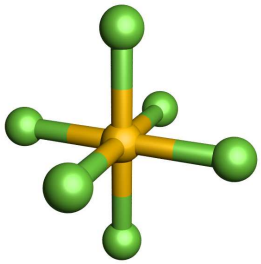


Figure 4.3: Experimental and theoretical stick spectra of the $^2B_1 \rightarrow {}^2A_2$ transition of the phenoxyl radical. The theoretical spectrum was generated from the transition assignments given with the FCHT [TA] spectrum using an interval of 8 cm^{-1} between two measurement points.

Initial state ($^2B_1$)				Final state ($^2A_2$)	
Mode	Harm.	Anh.[d]	Exp. [104]	Projection	Harm.
v'_1	180.6	184.1		v'_1	102.5
v'_2	374.7	371.4		v'_2	331.6
v'_3	450.4	443.4	446	0.86 v'_4	404.5
v'_4	455.2	474.2	472	v'_3	437.9
v'_5	534.6	524.3	520	0.77 v'_5 + 0.15 v'_7	521.7
v'_6	600.5	590.0	616	0.53 v'_7 + 0.22 v'_5	530.0
v'_7	617.6	641.3	635	v'_6	600.8
v'_8	764.9	785.0	784	0.62 v'_8	773.3
v'_9	800.5	788.0		0.69 v'_{10} + 0.21 v'_{11}	809.6
v'_{10}	806.3	795.3		v'_9	820.3
v'_{11}	913.3	905.6		0.34 v'_{11} + 0.28 v'_{10} + 0.25 v'_8	832.5
v'_{12}	953.1	973.2		v'_{13}	962.0
v'_{13}	970.9	963.9	977	0.84 v'_{12}	976.2
v'_{14}	989.4	976.1	1016	v'_{14}	980.8
v'_{15}	1011.5	994.0	1038	v'_{16}	1023.8
v'_{16}	1094.2	1075.4	1072	v'_{15}	1043.2
v'_{17}	1170.7	1152.0	1140	0.74 v'_{17} + 0.18 v'_{18}	1181.7
v'_{18}	1171.0	1152.8	1167	0.71 v'_{18} + 0.17 v'_{17}	1188.6
v'_{19}	1276.9	1256.7	1266	v'_{19}	1254.2
v'_{20}	1340.9	1323.3	1318	0.75 v'_{20}	1366.4
v'_{21}	1418.8	1395.3	1397	0.68 v'_{21} + 0.19 v'_{23}	1421.0
v'_{22}	1447.5	1415.3	1441	0.59 v'_{24} + 0.37 v'_{22}	1442.4
v'_{23}	1478.5	1462.8	1481	0.81 v'_{25} + 0.16 v'_{21}	1573.9
v'_{24}	1548.7	1515.8	1515	0.80 v'_{23} + 0.14 v'_{21}	1597.7
v'_{25}	1587.8	1562.6	1550	0.46 v'_{22} + 0.32 v'_{24}	1615.9
v'_{26}	3173.3	3046.4	3018	0.88 v'_{26}	3172.2
v'_{27}	3179.2	3058.7	3054	v'_{27}	3174.9
v'_{28}	3194.3	3071.9	3065	v'_{29}	3194.0
v'_{29}	3202.1	3082.0	3074	0.76 v'_{30} + 0.23 v'_{28}	3194.9
v'_{30}	3205.0	3083.0	3094	0.66 v'_{28} + 0.23 v'_{30}	3224.8

Table 4.2: Calculated and experimental frequencies of the initial ($^2B_1$) and final ($^2A_2$) states of the phenoxyl. The projection of the normal coordinates of the final state Q''_i on the basis set of the normal coordinates of the initial state Q'_j. The coefficients correspond to the squared elements of the Duschinsky matrix (J^2_{ij}). Only the relevant terms (with coefficients above 0.1) are shown and coefficients above 0.9 are rounded to 1.

[d] Anharmonic frequencies have been calculated with the N07 basis set which provides very good results for the anharmonic calculations [107]



4.3 Sulfur hexafluoride

4.3.1 Introduction

The sulfur hexafluoride (SF_6) finds a wide range of applications in industry, especially as an insulating gas. It has the ability to easily capture electrons, forming an anion with a long lifetime with respect to autodetachment. As a result, it is used for example to separate the winding in an electrical transducer from the magnetic circuit. This molecule has also a greenhouse effect and is one of the six main gases of this kind targeted by the Kyoto protocol to reduce their usage.

While the neutral form has been extensively studied (see reference [108] and references therein), the anion form (SF_6^-) is less known. Recently, a vibrationally resolved electron photodetachment spectrum has been published [109]. In this short study, we mainly focused on the working of our procedure and its ability to reproduce this spectrum.

The sulfur hexafluoride (SF_6^-) represents an interesting case of a relatively small and highly symmetric system where accurate *ab initio* methodologies can be applied, giving possibility for an easy computation of accurate theoretical spectra which can be of great value for the interpretation of the best available experimental spectroscopic data in the gas phase.

4.3.2 Computational details

For SF_6 and its negative ion, frequencies and geometries have been calculated at MP2 [110] level with aug-cc-pVTZ basis set [111, 112]. Moreover the electronic energies of the initial and final electronic states have been refined at the Coupled Cluster level [113–115]. The CCSD(T) calculations have been performed with the MOLPRO [116] package.

4.3.3 Discussion

The photodetachment spectrum of SF_6^- is characterized by a broad progression, with the band maximum shifted by more than 2 eV from the 0-0 transition corresponding to the adiabatic electron affinity (AEA). These features have been attributed to the significant changes in the geometry between ionic and neutral species [108, 109]. Indeed MP2/aug-cc-pVTZ calculations show a significant elongation of the S-F bond upon electron attachment (from 1.5750 to 1.7146 Å), while the molecule octahedral symmetry remains unchanged. Moreover, the theoretical [CCSD(T)/aug-cc-pVTZ] AEA of 1.06 eV is in very good agreement with the experimental value of 1.0 eV [117]. The fully theoretical photodetachment spectrum, calculated in an energy range from 0 to 5.0 eV within the Franck-Condon approximation on the basis of the aforementioned *ab initio* results (see Figure 4.4), clearly resembles its recently measured experimental counterpart [108].

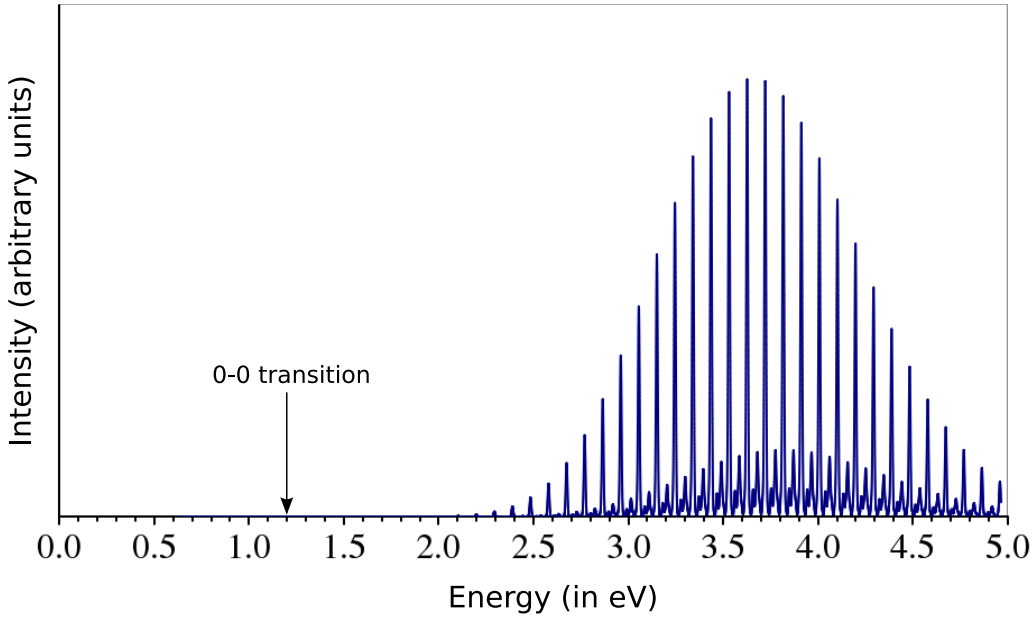


Figure 4.4: Theoretical electron photodetachment spectrum of SF_6^-. Full spectrum in a range from 0 to 5 eV calculated within Franck-Condon approximation with combinations between all modes considered; the energy of 0-0 transition is marked by an arrow.

For the SF_6^- photodetachment spectrum most of the vibrational progression derives from excitation of the single totally symmetric S-F stretching mode (v_1). Indeed, the computed spectrum shows a regular pattern of bands, the most intense corresponding to v_1.

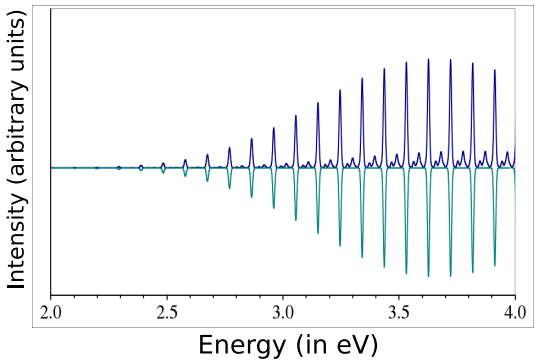


Figure 4.5: Comparison between the full spectrum (upper panel) and a spectrum with coupling between modes excluded (lower panel)

Figure 4.5 shows a comparison between the spectrum obtained with only the FC integrals from class $\mathcal{C}_1$ and the complete spectrum, where couplings between all modes are also taken into account. In the latter case it is found that the weaker bands gain intensity from excitation of this totally symmetric mode, being related to the combinations between the overtones of v_1 and doubly excited degenerate mode v_4 or v_5.

As a consequence of the large changes in the S-F bond length, the most intense transitions are related to high overtones, which must therefore be considered to reproduce accurately the spectrum features. Hence, the default parameter for the maximum number of quanta, `MAXC1`, is insufficient to cover all the spectrum. Indeed in the latter case only about 30 % of the spectrum intensity has been achieved in comparison to 94% when all necessary excitations have been taken into account. The comparison of the spectra calculated with `MAXC1` set to 100 and to 20 is shown in figure 4.6.

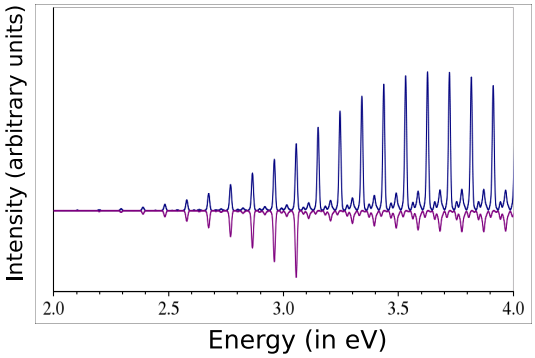
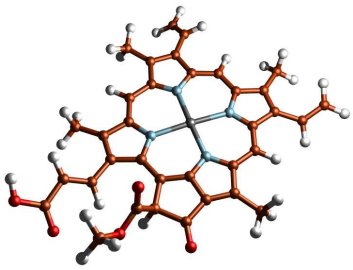


Figure 4.6: Comparison between the full spectrum calculated with $\mathcal{C}_{1_{\max}}$ set to 100 (upper panel) and the one with the original setting at $\mathcal{C}_{1_{\max}} = 20$ (lower panel)



4.4 Chlorophyll *c2*

4.4.1 Introduction

Despite ongoing experimental and theoretical research, the understanding of the molecular mechanism of light harvesting in photosystem II is not yet satisfactory. Quantum chemical computations of optical properties combined with spectroscopic experiments can undoubtedly contribute to shed further light on this phenomenon [118]. The triplet states of chlorophylls are of particular interest due to their dual photodamage and photoprotective role in photosystem II [119].

Chlorophyll *c2* (see Figure 4.7) is a large molecule with 73 atoms and 213 normal modes. Chlorophylls *c* can be found in some brown algae such as *fucus* or *diatoms* (a group of *eukaryotic algae*). Contrary to the more common chlorophyll *a*, they do not have the long *phytol* chain. Hence, being more compact around the central porphyrin structure, they are less flexible, making them easier to handle in the framework of the Franck-Condon principle.

The $T_1 \rightarrow S_0$ phosphorescence spectrum of Chlorophyll *c2* has been chosen to demonstrate various aspects related to spectrum convergence and applicability of the integrated approach[a]. In this study, we will focus on the spectrum convergence and especially on its meaning, depending on whether one wishes to reach the analytic limit or obtain the overall spectrum shape.

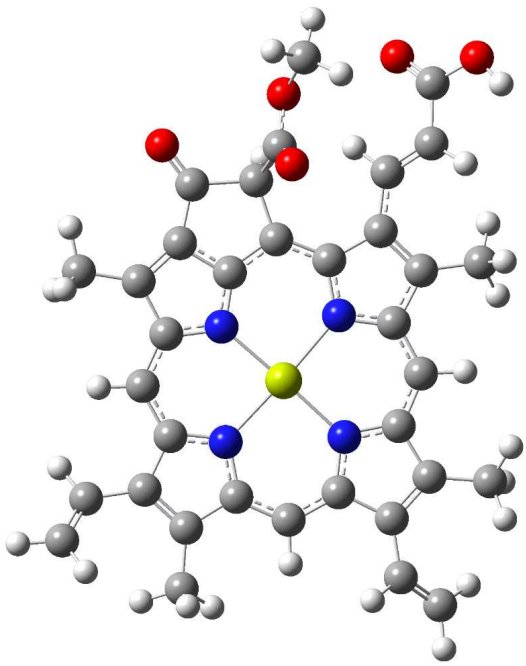


Figure 4.7: Structure of chlorophyll *c2*.

4.4.2 Computational details

Frequencies and geometries have been calculated, for both singlet ground and triplet excited electronic states, at the DFT level with the PBE0 [120] functional and 6-31G(d) basis set.

4.4.3 Discussion

When considering large systems such as chlorophyll *c2*, the maximum number of integrals to compute for each class, $N_I^{\max}$, can become a serious limitation. A reason for this is that the number

[a]The discussion presented here is centered on the procedure. A more thorough study is planned in the future

of combinations of the excited modes in the final state increase in each class with the number of normal modes following the relation

$$\frac{N!}{(N-n)!n!}$$

where n represents the class. We recall here that our prescreening method, described in section 2.8, chooses the maximum number of quanta $\boldsymbol{v}''_{\max}$ for each mode so that the number of transitions to compute does not exceed the limit $N_I^{\max}$. This number of transition is roughly evaluated as the product of the arithmetic mean $\langle \boldsymbol{v}''_{\max} \rangle$ and the number of combinations of the excited modes. Consequently, for the same value of $N_I^{\max}$, the larger is a system (an so N) and the lower will be $\langle \boldsymbol{v}''_{\max} \rangle$, reducing the quality of the calculations. A solution would be to increase $N_I^{\max}$ when dealing with larger systems. However, a higher number of transitions to compute in each state obviously improves the spectrum quality but at the same time strongly increases the required computational times and memory usage. To avoid as much as possible such an increase, it is necessary for our prescreening to be really efficient and choose correctly the most probable transitions.

Firstly, we will study the spectrum convergence with respect to the analytic sum.

Figure 4.8 shows the spectrum convergence with the increase of $N_I^{\max}$ from 10^2 to 10^{12}. As expected, a very small number of integrals is not sufficient, and leads only to about half of the spectrum intensity. Moreover, even as many as 10^{12} integrals cannot provide the full convergence of the spectrum intensity. It could be possible to increase further $N_I^{\max}$ but the computational costs are highly expensive at this point. However, as we will point out below, the expected gain is low and brings very little improvements to the analysis of the theoretical spectrum.

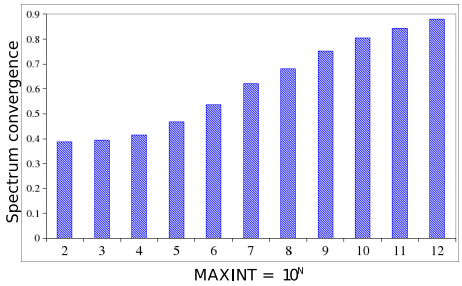


Figure 4.8: Convergence of the spectrum calculation for chlorophyll *c2* with the maximum number of integrals `MAXINT` set for each *class*

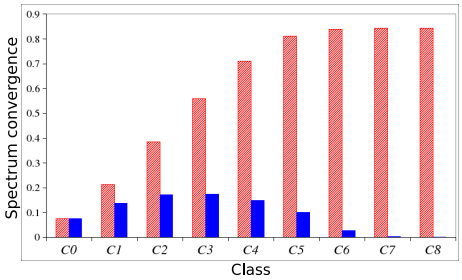


Figure 4.9: Convergence of the spectrum calculation for chlorophyll *c2* with *classes*. The total convergence up to class $\mathcal{C}_n$ is shown as red blocks, while contribution of class $\mathcal{C}_n$ is shown in blue.

Figure 4.9 shows the spectrum convergence with *classes* calculated with $N_I^{\max}$ set to 10^{11}. It is apparent that the contribution of *classes* higher than C_5 decreases steeply, and that the difference in spectrum intensity calculated up to $\mathcal{C}_7$ and up to $\mathcal{C}_6$ is smaller than 1 % (and below 0.1 % between $\mathcal{C}_7$ and $\mathcal{C}_8$), confirming the spectrum convergence with respect to *classes*. This shows that even for a large system such as the chlorophyll *c2*, it is unnecessary to treat a large number of classes. The default value `MAXBANDS=7` is adequately set and our discussions in chapter 3 assuming a maximum class $\mathcal{C}_{10}$ seem to cover a really large panel of molecular systems.

Nevertheless, in most cases the convergence of the spectrum line-shape is much faster [82] than the full convergence of the total spectrum intensity. This fact is particularly encouraging for large systems like chlorophyll *c2*. Figure 4.10 compares spectrum line-shapes calculated with $N_I^{\max}$ set to 10^2, 10^6 and 10^9. It is clear that the main spectral features are well reproduced even if total spectrum intensity is far from convergence. The spectra calculated with $N_I^{\max} = 10^9$ or larger are identical on this scale. Thus, inspection of the spectrum line-shape indicates that the most important transitions have been taken into account, and that reliable spectra have been computed already with $N_I^{\max}$ set to 10^9. The analysis of *classes* contributions to the total spectrum (Figure 4.11) shows that most of the spectrum bands are composed from *classes* up to $\mathcal{C}_4$, with $\mathcal{C}_1$ and $\mathcal{C}_2$ influencing most the spectrum line-shape.

Contributions of the classes related to the simultaneous excitation of five and more modes are much flatter, and of little importance for the spectrum line-shape, although they are not negligible for the spectrum intensity. The present case of chlorophyll turns out to be much more challenging than that adopted as a benchmark by Dierksen *et al.* [70] and Jankowiak *et al.* [83], a very large polycyclic aromatic hydrocarbon (PAH) derivative with 462 normal modes. Such PAH has a rather narrow photoelectron spectrum and our method is able to converge it up to values larger than 0.9. These tests show that our methodology can satisfactorily compute converged spectra also for large challenging systems. When the interest is focused on the high-energy wing of the spectrum (the one suffering of the largest relative error) as for instance for computation of nonradiative transition rates, a careful check of convergence in that energy region must be performed and purposely tailored methods may result more suitable.

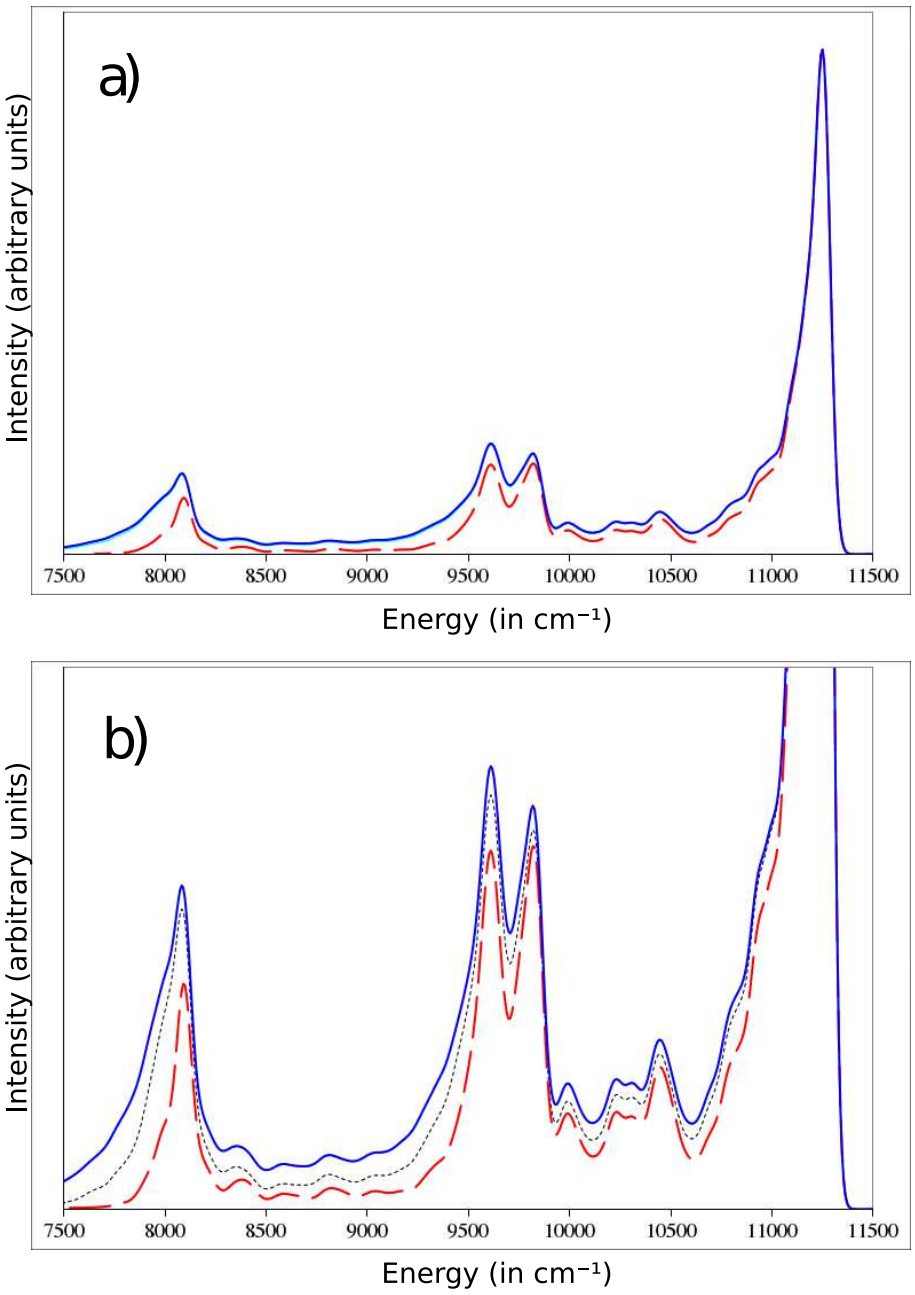


Figure 4.10: Convergence of the spectrum calculation for chlorophyll *c2* with the threshold on the number of computed integrals. Comparison of spectrum shape calculated with $N_I^{\max}$ set to 10^2 (dashed red line) and 10^9 (solid blue line) is shown on upper panel, while the onset with spectra calculated with $N_I^{\max}$ set to 10^2, 10^6 (fine-dashed black line) and 10^9 is shown on lower panel.

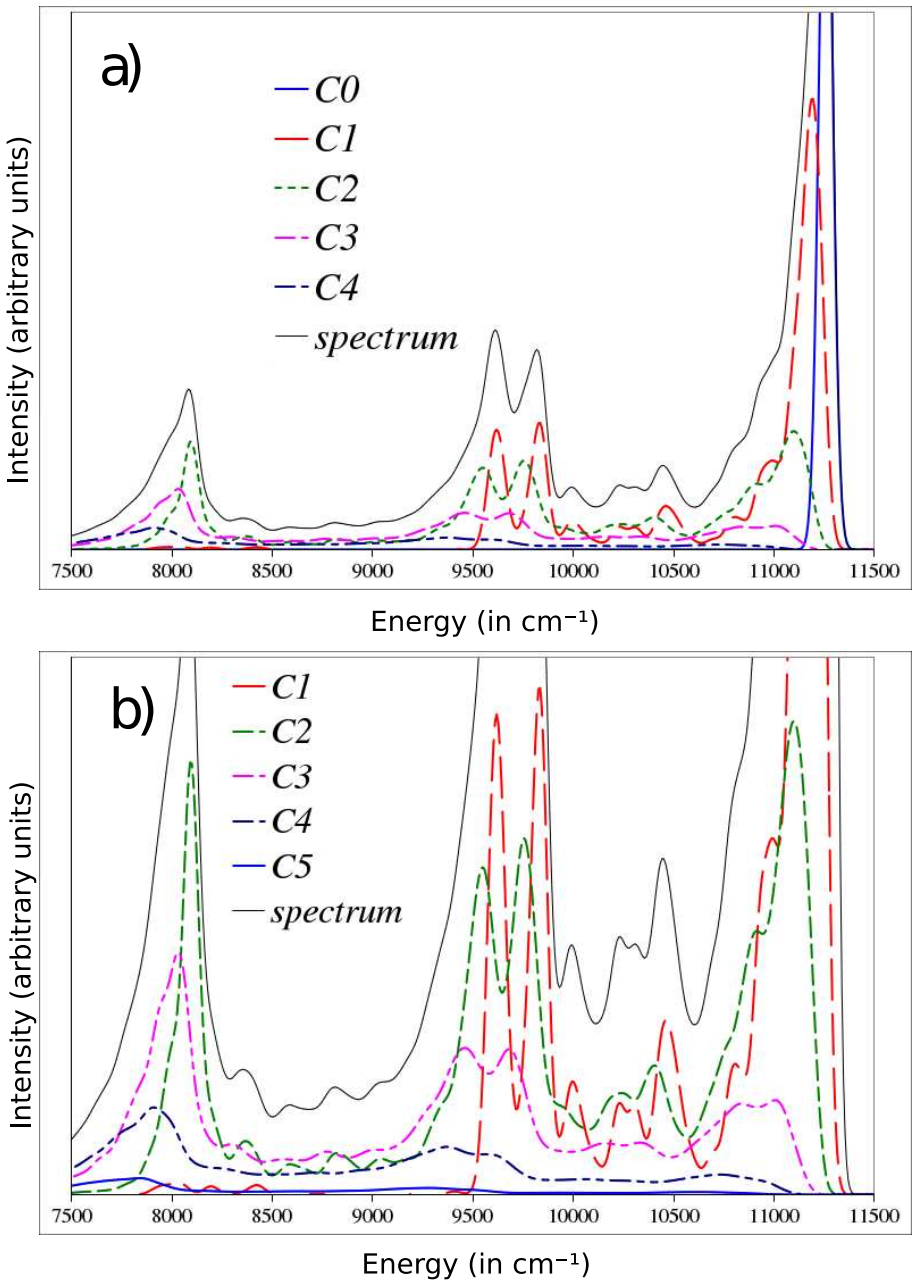
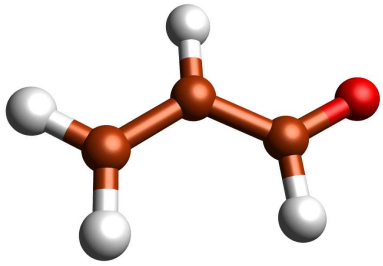


Figure 4.11: Convergence of the spectrum calculation for chlorophyll *c2* with *classes*; contributions of specific *classes* are compared with total spectrum (see legend). *Classes* $\mathcal{C}_0$ to $\mathcal{C}_4$ are shown on upper panel, while the onset with contribution of *classes* $\mathcal{C}_1$ to $\mathcal{C}_5$ is shown on lower panel. Contribution of higher *classes* are not visible in this scale.



4.5 Acrolein

4.5.1 Introduction

Acrolein is a relatively unstable compounds which is used in plastic and perfurme industries. It has several interesting chemical properties due to its aldehyde group and double bond. Also, because of their proximity, the double bond is activated by the aldehyde group, acrolein can play the role of a dienophile. As a consequence of these characteristics, the molecule finds numerous applications in organic syntheses.

The UV absorption spectrum of acrolein has attracted significant attention since this molecule exhibits two conjugated chromophores C=C and C=O, a common feature for many natural systems. In particular, a blueshift of the $n \rightarrow \pi^*$ transition of the C=O group has been observed in going from gas phase to aqueous solution (see Ref. [121] and references therein).

The procedure FRANCK is directly integrated inside GAUSSIAN and so, allows a straightforward computation of the gas phase and aqueous solution absorption spectra of acrolein, giving direct insights into the experimentally observed effect.

4.5.2 Computational details

The structures and frequencies of acrolein have been determined by DFT/TD-DFT computations with the B3LYP functional and N07 polarized double-ζ basis set [107], both in gas phase and in aqueous solution. The effect of water solvent has been included by means of the polarizable continuum model, where the solvent is represented by a homogeneous dielectric polarized by the solute, placed within a cavity built as an envelope of spheres centered on solute atoms [122]. The solvent is described in the non-equilibrium limit where only its fast (electronic) degrees of freedom are equilibrated with the excited-state charge density while the slow (nuclear) degrees of freedom remain equilibrated with the ground state. This assumption is sufficient to describe absorption spectrum in solution, due to the different time scales of the electronic and nuclear response components of the solvent reaction field [82].

4.5.3 Discussion

To simulate the spectrum line-shape it is necessary to convolute the stick-spectrum with a Gaussian with an appropriate full-width at the half maximum (FWHM): Figure 4.12 compares spectra calculated with the values of the FWHM set to 500 cm^{-1} and 1000 cm^{-1}. For acrolein the latter choice better reproduces the broad structure of the experimental spectrum. In the present approach it is possible to improve the Franck-Condon spectrum by considering changes of the transition dipole moment with the geometry. It is worth mentioning that in the present case inclusion of the Herzberg-Teller term does not require any additional quantum mechanical computation, since

the TD-DFT frequencies are calculated numerically giving direct access to the necessary derivatives of the transition dipole moment with respect to the normal coordinates of the excited electronic state. Inclusion of the HT term is particularly important for dipole-forbidden or weakly-allowed transitions where the FC approximation is unreliable.

This is the case of the weakly dipole allowed $n \rightarrow \pi^*$ transition of acrolein ($\mu = 0.0463$ a.u.), where the HT contribution indeed influences significantly the spectrum line-shape, as shown by the comparison of the FC and Franck-Condon Herzberg-Teller (FC-HT) spectra on Figure 4.13. Both the FC and the FC-HT spectra are fully converged (100.0 %) to their respective limits (see Section 2.9).

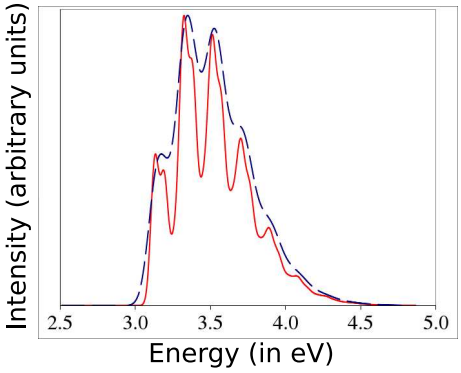


Figure 4.12: Theoretical absorption UV spectra of $n \rightarrow \pi^*$ electronic transition of acrolein: gas phase spectrum in a range from 2.5 to 5 eV calculated within Franck-Condon approximation with FWHM of 500 (red line) and 1000 cm^{-1} (blue dashed line).

The FC-HT spectra calculated in gas phase and in aqueous solution are compared in Figure 4.14. It is evident that not only the solvent shift is well reproduced by the theory, but that also changes in the band shapes agree with the recent results obtained with a more accurate but computationally demanding time-domain approach [121]. Moreover, the present method works directly in the frequency domain, so that all the individual vibronic contributions to the total spectrum are computed and can be easily assigned, as shown in Figure 4.15.

Figure 4.15 compares the convoluted spectra calculated within harmonic and anharmonic [TA] approximation (as described in section 3.3.1). The stick spectrum shows that the most pronunced progressions are related to only a few vibrations. On the whole, a good agreement is observed between the simulated stick spectrum and the experimental high reslution data stemming from supersonic free-jet expansion cavity ring-down spectroscopy [123]. The most intense bands have been assigned to the fundamentals or overtones of pure *C=C stretching* and *C=O stretching* [123] vibrations, while our results shows that a strong mixing prevails for these modes, as summarized in table 4.3. It should be mentioned that, although we obtained a good agreement between the computed anharmonic and experimental frequencies for the ground state, the theoretical results for the excited state are less accurate, showing a strong increase of anharmonicity upon an electronic excitation.

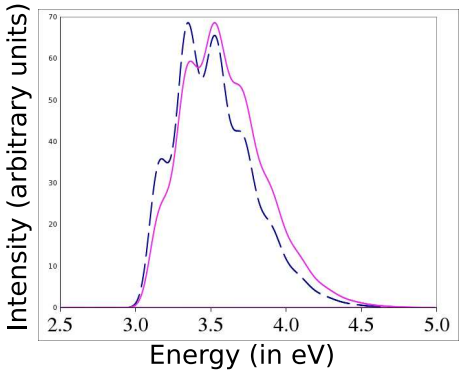


Figure 4.13: Theoretical absorption UV spectra of $n \rightarrow \pi^*$ electronic transition of acrolein: Comparison between gas phase spectra calculated within either the Franck-Condon (blue dashed line) or the Franck-Condon Herzberg-Teller approximations (pink line).

However, these results show the capacities of the implementation to correctly account for solvent effects in our vibrationally resolved electronic spectra.

We foresee that the accessibility to an ease and straightforward method for the computations of vibrationally resolved spectra within the integrated approach here described, may lead to breakthroughs in the studies of UV-vis spectra in condensed phases.

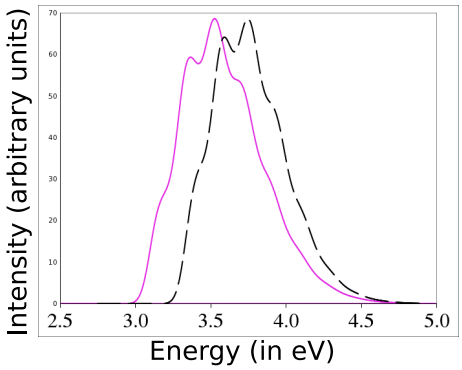


Figure 4.14: Theoretical absorption UV spectra of $n \rightarrow \pi^*$ electronic transition of acrolein: Comparison between calculated spectra for acrolein in gas phase (pink line) and in water solution described by the CPCM model (black dashed line).

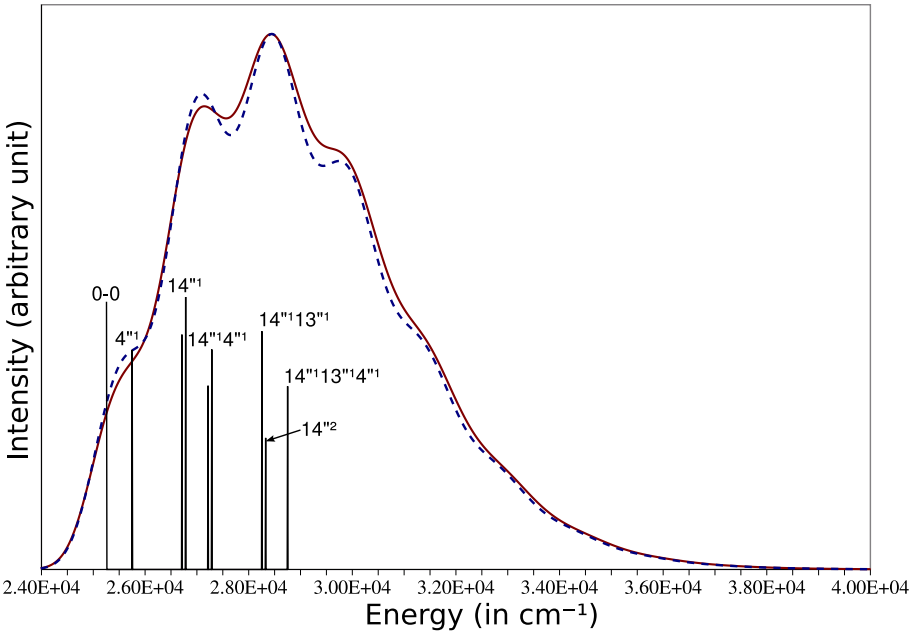
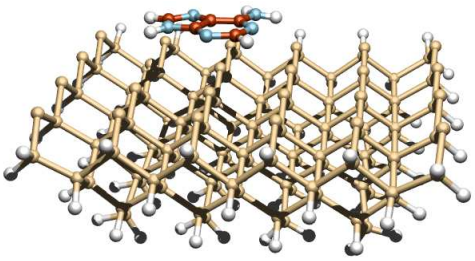


Figure 4.15: Assignment of the main bands of the theoretical absorption UV spectrum of $n \rightarrow \pi^*$ electronic transition of acrolein in gas phase. Red solid line reports the harmonic spectrum in a range from 24000 to 40000 cm^{-1} calculated within Franck-Condon Herzberg-Teller approximation with FWHM of 1000 cm^{-1} while the blue dashed line represents the anharmonic [TA] spectrum. The main stick bands are assigned as n^x where n is the excited normal mode and x its quantum number.

	State 0				State 1		
Mode	Harm.	Anh.	Exp. [124]	Assignment	Harm.	Anh.	Assignment
3	572	565	564	$v'_{\text{C=0 bending}}$			
4					507	501	$0.9v'_3$
12	1458	1421	1420	$v'_{\text{CH}_2\text{ bending}}$			
13	1681	1646	1625	$v'_{\text{C=C stretching}}$	1500	1468	$0.44v'_{12} + 0.21v'_{13} + 0.23v'_{14}$
14	1775	1746	1724	$v'_{\text{C=O stretching}}$	1568	1538	$0.51v'_{13} + 0.37v'_{14}$

Table 4.3: Selected vibrational frequencies of acrolein in its ground and excited electronic states. Computed and experimental frequencies (in cm^{-1}) are listed along with their assignment. The excited state normal modes are expressed as a linear combination of the ground state ones according to the Duschinsky rotation (only the highest elements are written), e.g $v''_j = \text{J}_{ji} v'_i$



4.6 Adenine on silicon (100)

4.6.1 Introduction

Biological molecules represent a natural source of elaborate nanostructures that could support the bottom-up construction of new materials of relevant technological interest. To fully grasp the possibilities they offer, it is necessary to have a good understanding of the influence of the nanoscale organization on the optical, chemical and electric properties of these materials. In particular, the study of the biomolecular systems in their excited electronic states is very important for the optics, photonics and sensoristics. Our approach is in line with such a demand as illustrated by simulation of the photoelectron spectrum of adenine adsorbed on Si(100) surface.

4.6.2 Computational Details

In the case of adenine molecule adsorbed on Si(100) surface, the ONIOM [125] QM/MM scheme has been adopted for the geometry structures and frequencies with the Si(100) surface represented by a cluster of 119 Silicon atoms. The QM part corresponding to the adenine molecule has been calculated at the B3LYP/6-31+G(d,p) level, while the cluster has been modeled by molecular mechanics using the UFF force field [126].

4.6.3 Discussion

The full valence photoelectron spectrum of adenine is composed from several overlapping excitations [127]. The present work is aimed to show feasibility of spectra simulations for nanosystems, thus only ionization from the highest occupied molecular orbital (HOMO) has been considered. The Si(100) surface has been modeled by a cluster of 119 silicon atoms, shown in Figure 4.16, resulting in a total system with 636 normal modes.

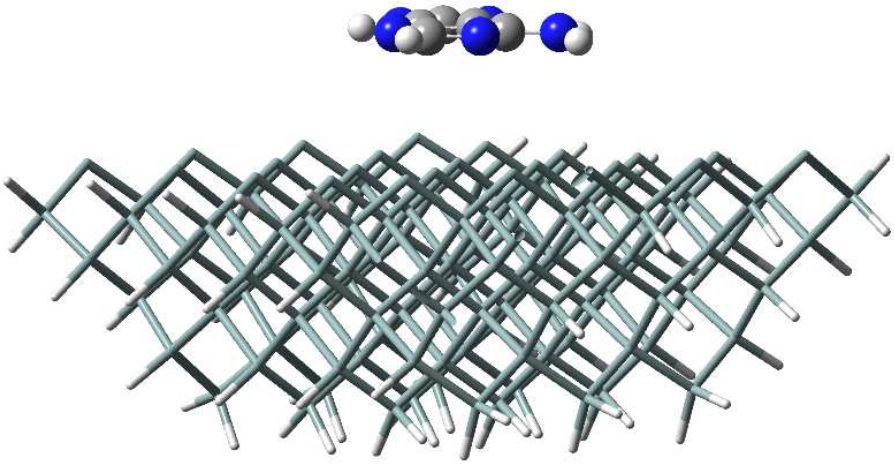


Figure 4.16: Adenine adsorbed on a cluster of 119 silicon atoms, modeling the Si(100) surface.

In order to put into evidence spectrum changes upon adsorption, the photoelectron spectra have been calculated for both isolated adenine molecule and adenine@Si(100), putting. Both spectra are plotted in the range of 8.0-8.7 eV roughly corresponding to the first band of valence shell photoelectron spectrum. Figure 4.17 shows the spectra in the absolute energy scale as well as in a relative scale where the 0-0 transition is set to zero. It can be seen that adsorption on Si surface yields a small red shift of the excitation origin, while new vibronic transitions corresponding to intermolecular vibrations modulate the spectrum line-shape.

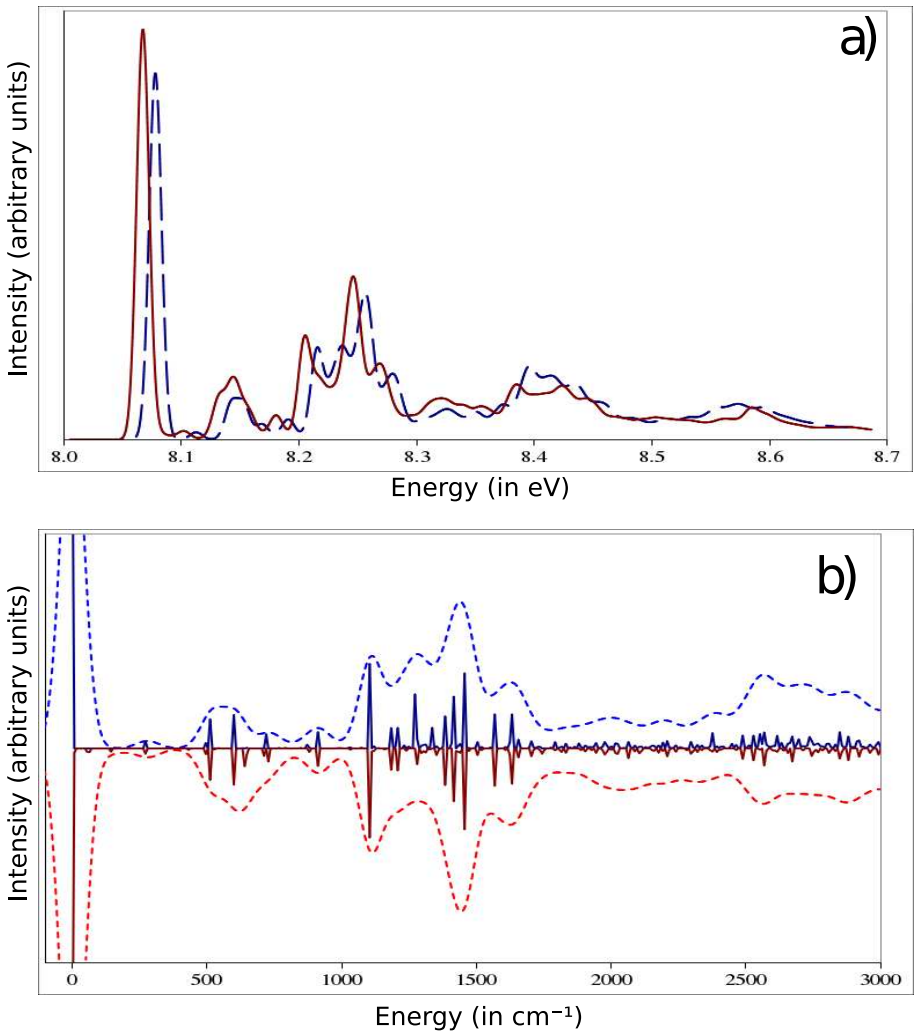


Figure 4.17: Comparison between the theoretical FC photoionization spectra in gas phase of isolated adenine (blue dashed line) and adenine adsorbed on a Si(100) surface (solid red line): a) spectra in an absolute energy range from 8.0 to 8.7 eV calculated within Franck-Condon approximation with FWHM=100 cm^{-1}; b) spectra shifted to the relative origins of the 0-0 electronic transitions, isolated molecule (upper panel) and adenine@Si(100) (lower panel); the stick bands show the most important transitions.

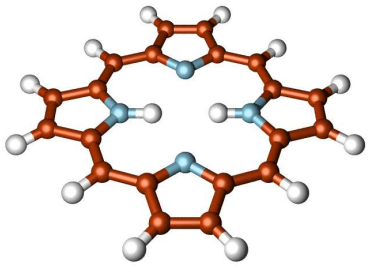
It is interesting to analyze the number of combinations for each class $\mathcal{C}_n$ for such a large system, which is directly related to the number of transitions to compute, and investigate the efficiency of the adopted selection procedure. Table 4.4 lists ${}_NC_n$ for isolated adenine and adenine@Si(100), and the spectrum intensity achieved with $N_I^{\max}$ set to the default value 10^8. It is noteworthy that in both cases, either an isolated molecule with 39 normal modes or a macrosystem with over 600,

	Adenine		Adenine@Si(100)	
Class (n)	${}_NC_n$	progression	${}_NC_n$	progression
3	9.14E+03	84.54%	4.27E+07	87.31%
4	8.23E+04	93.57%	6.75E+09	94.82%
5	5.76E+05	97.48%	8.54E+11	97.37%
6	3.26E+06	98.32%	8.98E+13	97.88%
7	1.54E+07	98.39%	8.08E+15	97.93%

Table 4.4: Convergence of spectra computations for adenine and adenine@Si(100). For each class $\mathcal{C}_n$ the number of combinations of the n excited oscillators ${}_NC_n$ and spectrum progression are listed. The $\mathcal{C}_1$ and $\mathcal{C}_2$ transitions have been computed by analytical formulae allowing a maximum quantum number $\mathcal{C}_{1_{\max}}$=30, and $\mathcal{C}_{2_{\max}}$=20 (`MAXC1`=30, `MAXC2`=20) respectively. For the *classes* $\mathcal{C}_n$, $n \geq 3$ the transitions to be computed have been selected setting the parameter $N_I^{\max}$ to 10^8 (the default value).

almost all spectrum intensity (about 98 %) has been recovered at an equivalent computational cost.

For the cluster, the default value of `MAXINT` is not sufficient to consider the whole initial pool even for only three simultaneously excited modes ($\mathcal{C}_3$ class). This particular case shows the ability of the *a priori* strategy to select only the relevant transitions and discard the less probable ones.



4.7 Porphyrin

4.7.1 Introduction

We present here a last molecular system that can be linked to the previously described chlorophyll *c2*. Porphyrins are very important chromofores, involved in many relevant biological processes like photosynthesis. Free base porphyrin (H2P), sketched in figure 4.18, is similar to the major building block (Mg-Porphyrin) of chlorophyll and it is therefore very interesting to study its absorption of UV and visible light and the corresponding emission properties.

These are nicely interpreted according to the four-orbitals model of Gouterman [128]. For an idealized D_{4h} symmetry (actually shown by some metal-porphyrin complexes) two pairs of degenerate states give rise to the two bands Q (dark) and B (intense, also called Soret band). Each of these bands is actually composed by a couple of transitions, namely (Q_x,Q_y), and (B_x,B_y) with components polarized in the x and y directions (considering the molecule in the xy plane oriented according to figure 4.18). H2P actually belongs to the D_{2h} point group and the degeneracy between the Q_x, Q_y and the B_x, B_y transitions is removed. However the Q bands are still very weak and a proper calculation of their spectra must take into account the Herzberg-Teller effect. Recently, a full coordinate analysis of the absorption spectrum of H2P has been published [129], which includes HT effect but relies on the previously discussed LCM approximation [11], thus neglecting both Duschinsky couplings and the changes in the oscillation frequencies of the normal modes in the two electronic states. However, the simultaneous treatment of Duschinsky and Herzberg-Teller couplings is necessary to account exactly (in the harmonic approximation) for the role of nuclear vibrations in an optical transition.

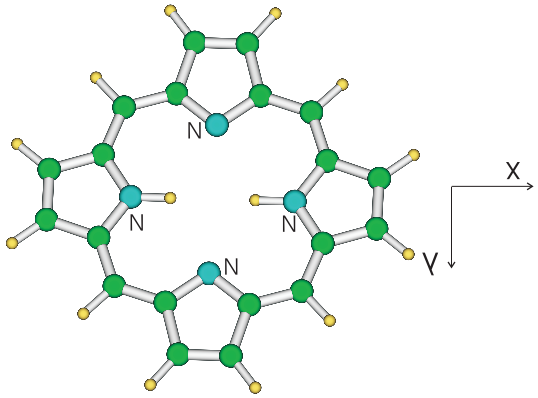


Figure 4.18: Schematic drawing of free base porphyrin (H2P). The molecule is in the xy plane and the x axis passes through the opposite hydrogenated N atoms.

The approach used here differs slightly from the implementation in GAUSSIAN described previously. The reason is that the study was done with a standalone version of the prescreening, FCCLASSES [95], and some new features were applied to it. This example remains perfectly feasible in the procedure albeit with some minor changes and shows for the first the first time, to the best of our knowledge, such an exact calculation for the absorption and fluorescence spectra of the Q_x band of H2P.

4.7.2 Computational details

The equilibrium structures and vibrational frequencies of the ground and first-excited state of H2P have been obtained in the frame of the DFT and its time-dependent (TD-DFT) extension for the excited state, employing the hybrid functional PBE0 [120] and the standard 6-31G(d) basis set. The derivatives of the electronic transition dipole moment ($\boldsymbol{\mu}_{if}$) with respect to the normal coordinates are obtained, at no additional computational cost, as a side-product of the numerical computation of the second derivatives of the TD-DFT excited state energy, needed for the harmonic analysis at the equilibrium geometry of the excited-state. To generate both absorption and emission spectra and be able to confront them, a consistent choice for the state of reference to perform the Taylor expansion of $\boldsymbol{\mu}_{if}$ is necessary. As explained previously (see the conclusion of chapter 3), the best choice would be the initial state. As presented in the description of FRANCK, the current implementation works with the final state as the reference for the Taylor expansion. While we presented in conclusion of our implementation a possible treatment with respect to the initial state, the method used here proceeds differently, assuming that the electronic transition dipole moment can be accurately described with a sufficiently developed Taylor series. We assumed here that the Herzberg-Teller approximation gives a very good approximation of $\boldsymbol{\mu}_{if}$. In this case, it is possible to write the following relation:

$$\boldsymbol{\mu}_{if} = \boldsymbol{\mu}_{if}(\mathbf{Q}''_0) + \sum_{k=1}^{N} \left(\frac{\partial \boldsymbol{\mu}_{if}}{\partial Q''_k}\right)_0 Q''_k = \boldsymbol{\mu}_{if}(\mathbf{Q}'_0) + \sum_{k=1}^{N} \left(\frac{\partial \boldsymbol{\mu}_{if}}{\partial Q'_k}\right)_0 Q'_k \tag{4.1}$$

Using equation 2.51, it is possible to subsitute the normal coordinates of the final state $\mathbf{Q}''$ in the previous equality:

$$\boldsymbol{\mu}_{if} = \boldsymbol{\mu}_{if}(\mathbf{Q}''_0) + \sum_{k=1}^{N} \left(\frac{\partial \boldsymbol{\mu}_{if}}{\partial Q''_k}\right)_0 \sum_{l=1}^{N} \mathrm{J}_{kl}{}^{\mathrm{T}} Q'_l - \mathrm{J}_{kl}{}^{\mathrm{T}} \mathrm{K}_l = \boldsymbol{\mu}_{if}(\mathbf{Q}'_0) + \sum_{k=1}^{N} \left(\frac{\partial \boldsymbol{\mu}_{if}}{\partial Q'_k}\right)_0 Q'_k \tag{4.2}$$

Hence, we obtain an equivalence between $\boldsymbol{\mu}_{if}(\mathbf{Q}''_0)$ and $\boldsymbol{\mu}_{if}(\mathbf{Q}'_0)$:

$$\boldsymbol{\mu}_{if}(\mathbf{Q}'_0) = \boldsymbol{\mu}_{if}(\mathbf{Q}''_0) - \sum_{k=1}^{N} \left(\frac{\partial \boldsymbol{\mu}_{if}}{\partial Q''_k}\right)_0 \sum_{l=1}^{N} \mathrm{J}_{kl}{}^{\mathrm{T}} \mathrm{K}_l$$

It is straightforward to see that if the HT approximation is not good enough, then the previous relation is erroneous. This would mean that a higher degree of the Taylor series should be taken into accound. As a matter of fact, knowing $\boldsymbol{\mu}_{if}(\mathbf{Q}''_0)$ and $\boldsymbol{\mu}_{if}(\mathbf{Q}'_0)$, it is possible to check the relevance of the HT approximation. Such an approach is interesting when $\boldsymbol{\mu}_{if}(\mathbf{Q}'_0)$ is unknown or cannot be easily computed.

Here, the TD-DFT transition dipole moment $\boldsymbol{\mu}_{if}$ is polarized along x (see figure 4.18) and in atomic units its x-component is 0.01802 and -0.1017 at the equilibrium geometry of the ground and excited states, $S_0 - min$ and $S_1 - min$ respectively. Computing $\boldsymbol{\mu}_{if}(\mathbf{Q}'_0)$ with the previous relation (using $\boldsymbol{\mu}_{if}(\mathbf{Q}''_0)$), we obtain 0.02085 au, a value in very close agreement with the TD-DFT result, supporting the validity of the HT approximation for the present case.

As described in section 2.9, two quantities are defined to check the spectrum convergence, SP_{FC} for the FC contribution and SP_{FCHT} for the FCHT spectrum. The condition $SP_{\mathrm{FC}} = 1$ is necessary

but not sufficient to obtain $SP_{\mathrm{FCHT}} = 1$. It should be noted that the shape of the spectrum is expected to converge faster than these quantities do as shown in section 4.4 and in references [82,94].

4.7.3 Results and discussion

Absorption and fluorescence spectra

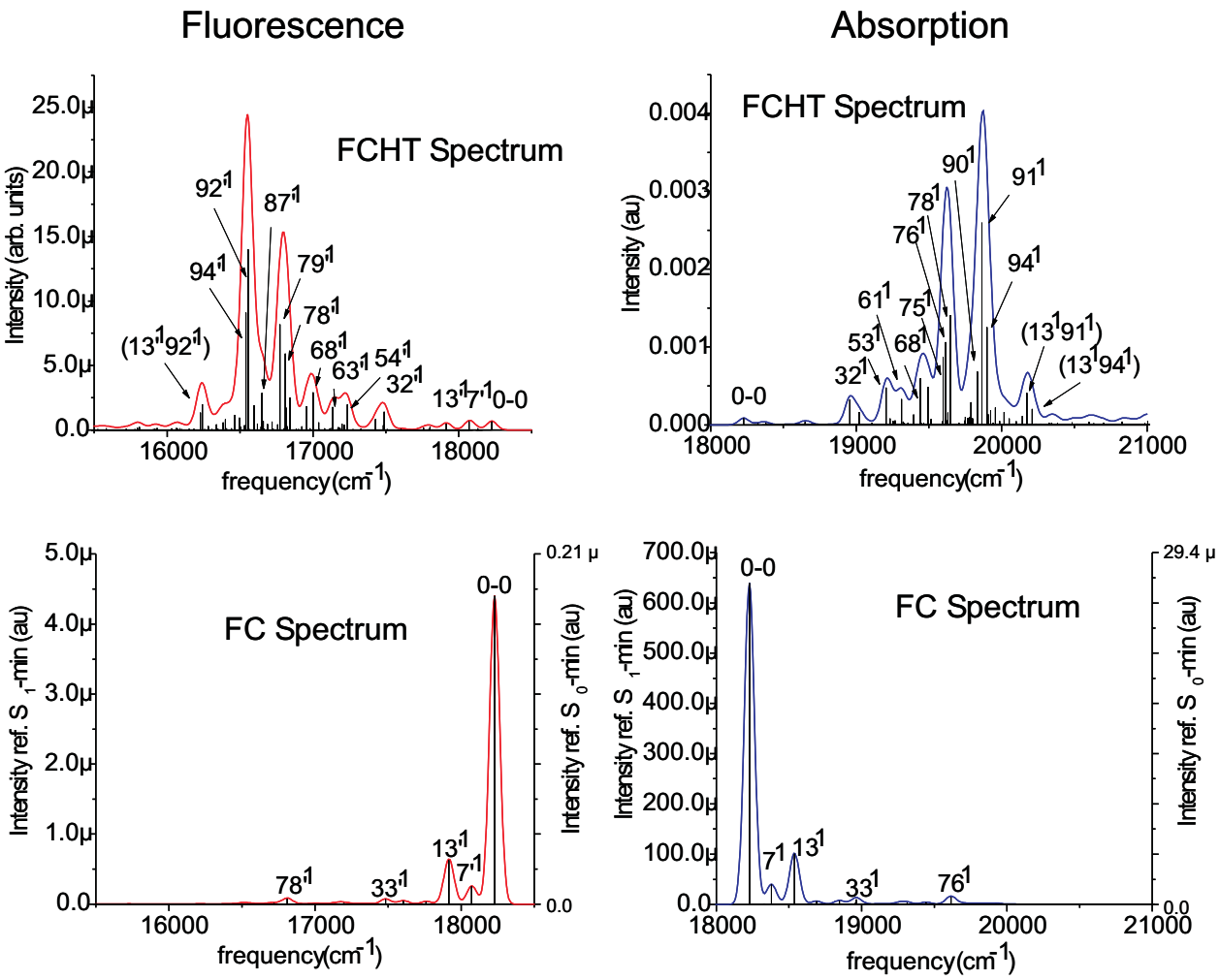


Figure 4.19: Fluorescence (left panels) and absorption (right panels) computed spectra of H2P including only the FC contribution or both FC and HT contributions (FCHT). For FC spectra the intensity is affected by the choice of the reference geometry, as shown by comparison of the left and right vertical axes. The intensities on the vertical axes are reported in a.u. (see text) and μ stands for 10^{-6}. The spectra, computed at T=0 K, have been convoluted with a gaussian with a FWHM=80 cm^{-1}. Stick spectra are also reported and the main bands have been assigned as n^x where n is the excited normal mode and x its quantum number. Combination bands are reported in parentheses.

In the upper panels of figure 4.19 we report the absorption and fluorescence spectra computed at T=0 K including both FC and HT contributions (FCHT) and convoluted with a gaussian with a full width at half maximum FWHM=80 cm^{-1}. The intensities are computed according to equations 1.20, for the absorption, and 1.21, for the emission, skipping the prefactor $4\pi^2/(3c)$ and expressing the remaining quantities in atomic units. The maximum number of transitions to compute per class, $N_I^{\max}$, is set to 10^7, giving for absorption $SP_{\mathrm{FC}} \approx 1$ and $SP_{\mathrm{FCHT}} \approx 1$.

State S_1			State S_0	
Mode	Sym	ω'' (cm^{-1})	Projection[a]	$\Delta\omega$[b] (cm^{-1})
v''_{7}	a_{1g}	151.89	v'_{7}	-4.11
v''_{13}	a_{1g}	310.08	v'_{13}	-2.23
v''_{32}	a_{1g}	729.79	0.86 v'_{32} + 0.14 v'_{33}	-10.61
v''_{33}	a_{1g}	735.06	0.86 v'_{33} + 0.14 v'_{32}	-9.42
v''_{41}	b_{1g}	793.99	v'_{41}	-4.25
v''_{53}	a_{1g}	977.93	0.99 v'_{54}	-13.17
v''_{61}	a_{1g}	1085.68	0.99 v'_{63}	-5.28
v''_{68}	a_{1g}	1213.84	0.98 v'_{68}	-12.86
v''_{70}	b_{1g}	1266.58	v'_{71}	-4.86
v''_{75}	b_{1g}	1369.50	0.92 v'_{75}	-15.48
v''_{76}	a_{1g}	1387.82	0.85 v'_{78} + 0.07 v'_{80}	-30.97
v''_{78}	b_{1g}	1420.59	0.97 v'_{79}	-32.27
v''_{90}	a_{1g}	1606.43	0.95 v'_{91}	-25.88
v''_{91}	b_{1g}	1637.01	0.99 v'_{92}	-35.82
v''_{94}	a_{1g}	1671.76	0.94 v'_{94}	-15.75

Table 4.5: Relevant normal modes of the two electronic states

The FCHT spectra can be compared with the FC ones reported in the lower panel of figure 4.19 and obtained neglecting the transition dipole derivatives (HT terms) in the Taylor series. As discussed in section 3.9, the intensity of the FC contribution may change sensibly by adopting different reference geometries. This is shown by the double vertical axes of the lower panels of figure 4.19 that report the FC intensities computed from the zero-order term of the $\boldsymbol{\mu}_{if}$ expansion about the equilibrium geometry of the initial state ($S_0 - min$) (right axes) or the final state ($S_1 - min$) (left axes).

The comparison of the FC and FCHT spectra shapes shows the dramatic impact of the HT terms on the spectra. The FC spectra are very narrow, and are by far dominated by the 0-0 transition (between vibrational ground states). This result depends on the stiffness of the molecule, which shows only small displacements between the equilibrium structures of the two electronic states. The other peaks are due to transitions to excited vibrational states of the final state and are labeled as n^x where n represents the excited normal mode v''_n and x its quantum number.

The spectrum intensity is mainly related to the progression of normal modes with a_{1g} or b_{1g} symmetry, which are reported in table 4.5. The a_{1g} modes give rise to progressions which have both FC and HT contributions (polarized along x), while b_{1g} modes progressions result only from HT contributions (polarized along y). Symmetry considerations show that HT terms polarized along x mainly reflect intensity borrowing from the B_x Soret band, while y-polarized HT terms derive from mixing with the B_y state (minor contributions may also come from the Q_y state).

The normal modes of the excited state which contribute most to the FCHT absorption spectrum are reported in order of increasing frequencies in the first column table 4.5. The second and third columns report respectively its symmetry and the frequency. The fourth column shows the modes of the ground state S_0 on which the mode v''_n is mainly projected. The coefficients of the combinations correspond to the squared elements of the Duschinsky matrix ($\mathrm{J}(v'_i, v''_n)^2$). Inspection of the data

[a] The coefficients correspond to the $\mathrm{J}(m'_i, n'')^2$ elements.

[b] $\Delta\omega = \omega''_i - \omega'_j$, where j is the mode v'_j of S_0 which has the highest projection coefficient with v''_i.

shows that the Duschinsky mixing is not very large at least for the most important modes in the spectrum, the most mixed mode of S_1 being $76''$ with a coefficient $\mathrm{J}(v'_{78}, v''_{76})^2$ equals to 0.85. The last column of table 4.5 reports the frequency difference between v''_n and the normal mode of the initial state with the highest projection coefficient v'_i ($\omega''_n - \omega'_i$).

The FC fluorescence and absorption spectra in figure 4.19 show a good mirror-symmetry. The highest peak corresponds to the 0-0 transition while the second highest one is due to the excitation of one quantum on the oscillator v''_{13} (in absorption) and v'_{13} (in emission), which are fully projected one into the other (see table 4.5). The FCHT spectra are very different from their FC counterparts, showing a much richer structure and being much more shifted toward the red (fluorescence) and toward the blue (absorption), because of the negligible intensity of the 0-0 transition. This is due to the fact that the HT contributions are much stronger than the FC ones.

The FCHT absorption and fluorescence spectra show some deviation from mirror-symmetry that are more evident when looking to the assignments of the main bands. Some of the differences are only seeming: for instance, the strongest stick band is $1''_{91}$[c] in absorption and $1'_{92}$ in emission, but table 4.5 shows that these two normal modes are physically the same, being projected one into the other with a weight 0.99 (the same happens for the bands $1''_{53}$ in absorption and the $1'_{54}$ in emission). On the contrary, some differences have a physical meaning: for example the third (in order of decreasing intensity) stick bands underlying the two highest absorption peaks (namely $1''_{90}$ and $1''_{75}$) have much weaker counterparts in emission as a consequence of the interplay of Duschinsky and HT effects. Relevant FC/HT interference effects are seen in the region close to the 0-0 transition-frequency and a separate paragraph will be devoted to their discussion.

The limited displacements and frequency shifts, together with the moderate Duschinsky mixing, explain why most of the main transitions in both the FCHT spectra actually correspond to the fundamental $0 \rightarrow 1$ transition of a single mode. Only in the region of the spectra more distant from the 0-0 transition (the "red" region in emission and the "blue" region in absorption) the intensities of some combination bands really dominates the shape of the spectrum.

Study of the FC/HT interferences close to the 0-0 band

Some years ago Hohlneicher and coworkers reported a very interesting analysis of the FC/HT interferences [130–132] on a series of organic molecules. They found out that often a constructive interference in absorption becomes destructive in emission and viceversa. This phenomenon was rationalized by showing that for $0 \rightarrow 1$ transitions, the FC transition amplitude is expected to change sign when considering absorption or emission, unlike the HT one, thus turning a constructive interference into a destructive one. This feature was already noticed by Small [133] and experimentally investigated in H2P by Gradyushko *et al.* [134, 135]. In H2P, FC/HT interferences can only take place close to the 0-0 transition, where FC contributions are not negligible. Therefore, we investigate in a deeper detail the region in the interval of 700 cm^{-1} from the 0-0 energy. Mirror-symmetry is expected between the squares of the transition dipole moment elements $\left(\left|\langle\, \boldsymbol{v}' \mid \boldsymbol{\mu}_{if} \mid \boldsymbol{v}'' \,\rangle\right|^2\right)$ for absorption and emission. These data are reported as stick bands in figure 4.20, where the 0-0 frequency is set to zero and the 0-0 intensities are set to 1.

Figure 4.20 puts into evidence a breakdown of the absorption/emission mirror-symmetry with a

[c] $1''_{91}$ represents 1 quantum for the mode v''_{91}, and corresponds here to the vibrational state $|\, \mathbf{0}'' + 1''_{91} \,\rangle$

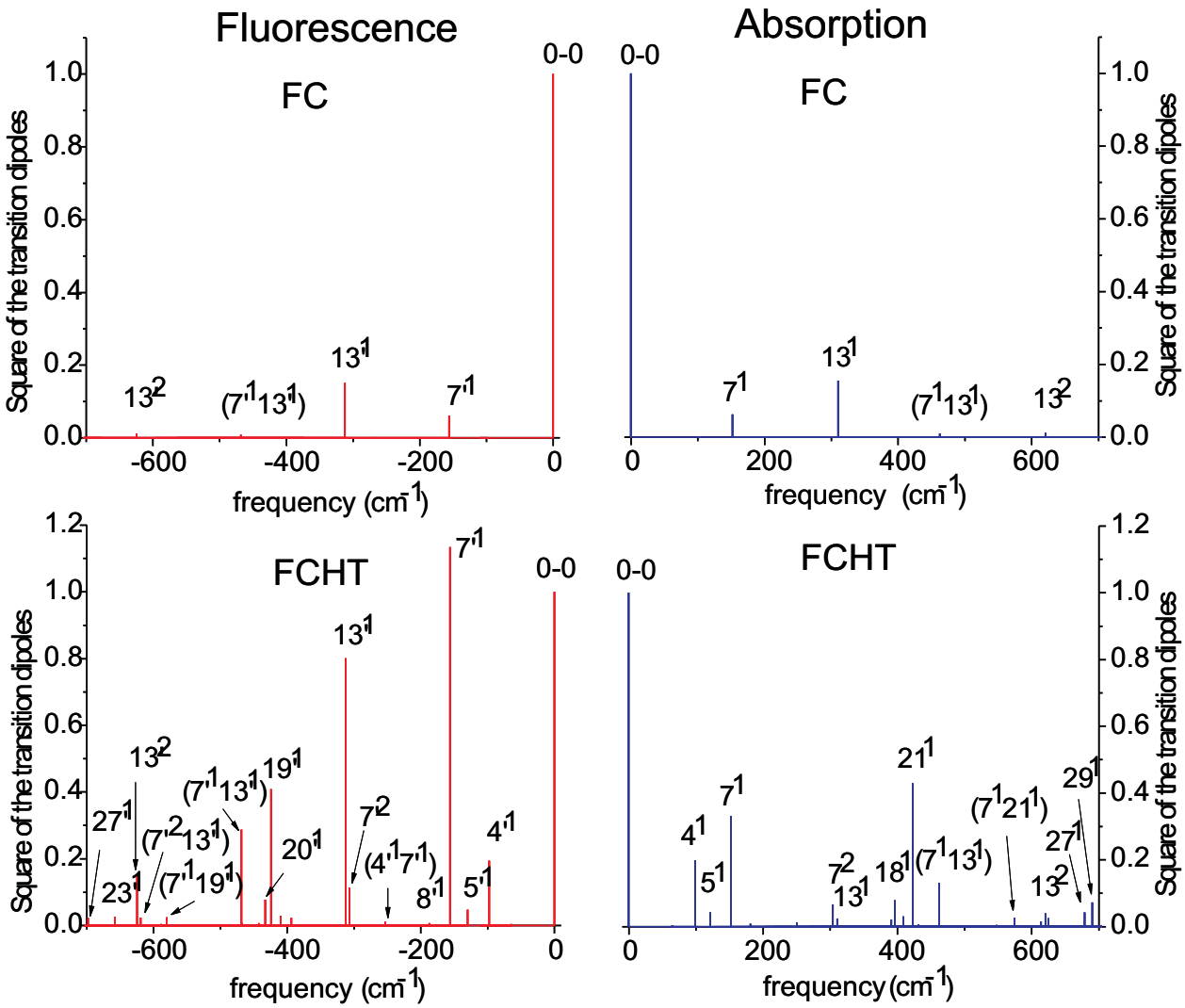


Figure 4.20: Squares of the transition dipole moments of the T=0 K FC and FCHT absorption and emission spectra, within a range of 700 cm^{-1} from the 0-0 transition frequency (set to zero), and normalized so that the 0-0 stick-band (the same in absorption an emission) has height 1.

huge enhancement of the emission bands $1'_7$ and $1'_{13}$ with respect to the absorption counterparts $1''_7$ and $1''_{13}$. This result cannot be due to Duschinsky effects since mode v''_7 and v'_7 coincide as well as v''_{13} and v'_{13} (see table 4.5). In order to assess if the mirror-symmetry breakdown is due to FC/HT interferences, we report in table 4.6 the values of the FCHT, FC and HT x-polarized transition amplitudes ($\langle \boldsymbol{v}' \mid \boldsymbol{\mu}_{if} \mid \boldsymbol{v}'' \rangle$) for these two bands, using as reference geometry for the Taylor expansion of $\boldsymbol{\mu}_{if}$ either $S_0 - min$ or $S_1 - min$. These two sets of data show that the interpretation of the observed enhancement in emission depends on the choice of the geometry reference, confirming the discussion in conclusion of chapter 3. For the $1''_7$ band, taking the $S_1 - min$ geometry as a reference leads to attributing to the FC/HT interference the cause of the mirror-symmetry breakdown, whereas the origin of the phenomenon is purely HT if considered with respect to $S_0 - min$. As for the $1''_{13}$ band, using $S_1 - min$ as reference geometry the small absorption intensity is a clear FC/HT destructive interference effect, while the remarkable emission intensity is purely due to the FC contribution. At variance, adopting the $S_0 - min$ reference the absorption intensity comes out from two very small FC and HT contributions, while the emission intensity is now almost a pure

	TD-DFT results			
	transition $1''_7$		transition $1''_{13}$	
abs FCHT	-0.0185		0.0050	
emi FCHT	-0.0343		0.0289	
	$S_0 - min$[d]	$S_1 - min$[d]	$S_0 - min$[d]	$S_1 - min$[d]
$\mu_{if}(x)$	0.02085	-0.1017	0.02085	-0.1017
abs FC	-0.0045	0.0219	0.0071	-0.0346
abs HT	-0.0140	-0.0404	-0.0021	0.0395
emi FC	0.0044	-0.0216	-0.0068	0.0340
emi HT	-0.0387	-0.0128	0.0358	-0.0052

Table 4.6: FC/HT Interferences for x-polarized transitions 1_7 and 1_{13}. For brevity only the notation for absorption bands is used. This creates no ambiguity since S_0 v'_7 and v'_{13} modes coincide with the S_1 v''_7 and v''_{13} ones, respectively.

HT effect. Finally, adopting the $S_0 - min$ reference for absorption and the $S_1 - min$ reference for emission, no sensible FC/HT interferences would be seen and one would ascribe the absorption/emission difference simply to a variation of the FC contribution as a result of the different values of $\mu_{if}(\mathbf{Q}_0, x)$ at the two reference geometries.

The reason why the same physical feature can have different interpretations is rather subtle: Hohlneicher *et al.* [130–132] noticed that in absence of strong Duschinsky couplings, once chosen a reference geometry, the FC contributions of fundamental transitions (1_n) are expected to be similar in magnitude but of opposite sign, while the HT contributions do not change sign. Nonetheless, the HT contributions may be very different in magnitude (see table 4.6) since, if the coordinate operator Q'' acts on a ground vibrational state (usually giving the smaller contribution) in absorption, it acts on a vibrational state with 1 quantum in emission (and viceversa). At variance, HT contributions in absorption and emission have similar magnitudes considering each of them in a different reference geometry. On the other hand, the magnitude of the FC contribution may be deeply different adopting different reference geometries, which clearly reflects in the relevance of the FC/HT interference term.

Effect of the Duschinsky mixing and frequency changes

In order to check the difference between our calculation and the results with a LCM-similar approach, we re-computed the FCHT absorption spectrum at T=0 K after switching off the Duschinsky couplings (setting $\mathbf{J} = \mathbf{I}$) and assigning to the normal modes on S_1 the same frequencies they have on S_0, as described in section 4.8.3. As previously explained, our method differs slightly from a pure LCM one since it takes into account the exact displacement of the atoms between the optimized structures of the initial and final state. Nevertheless, we will assume that it is an accurate description of the Linear Coupling Method and will refer to it as LCM.

Comparison of the LCM and FCHT spectra is shown in figure 4.21.

The largest difference is visible in the upper panel of the figure and corresponds to a global shift in the spectrum caused by the neglect of the difference between the S_1 and S_0 frequencies (the LCM 0-0 transition frequency is blue-shifted by $\approx$ 480 cm^{-1}).

[d]Reference geometry for the Taylor expansion of the transition dipole moment $\boldsymbol{\mu}_{if}$

In order to highlight other discrepancies, this difference is cancelled by setting in both cases the energy of the 0-0 transition to 0 cm^{-1}, hence switching to relative transition energies. The result is shown in the middle panel of figure 4.21 and the lowest panel of the figure represents an enlargement of the zone corresponding to the two highest peaks in the middle panel. The heights of all bands slightly change, and a sensible difference appears in the relative height of the two main bands. Additionally, the stick spectra shown in the lowest panel of figure 4.21 show a noticeable re-distribution of the intensities. However, the general shape of the spectrum changes only moderately, as it might have been expected on inspection of table 4.5, which shows that the mixing of the normal modes is weak and the changes of the frequencies are small. Because of that, our results substantially confirm the ones obtained for the absorption spectrum of H2P by Minaev *et al.* adopting the LCM method. [129]. On the other hand our calculations indicate that strongly mixed modes exist also in porphyrin. One example is the S_1 mode v''_{74} whose maximum projection on a single S_0 mode is 0.55. However these mixed modes are not active either by the FC or the HT mechanisms, and therefore they do not contribute significantly to the spectrum.

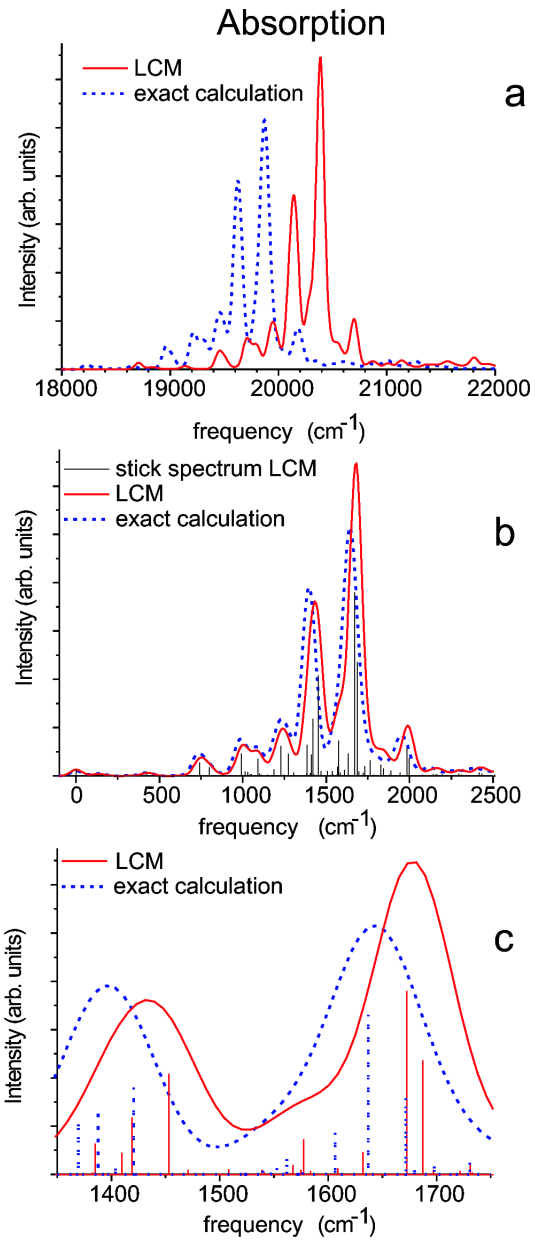
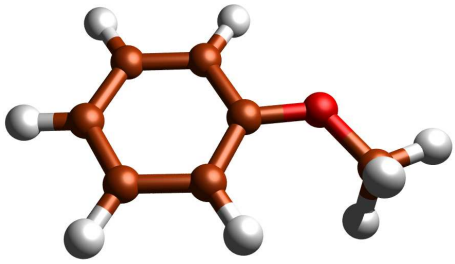


Figure 4.21: Upper panel: comparison of the T=0 K absorption spectrum of H2P obtained by an exact calculation and by only considering equilibrium position displacements (DO). The spectra have been convoluted with a gaussian with a FWHM=80 cm^{-1}. In the middle and lower panels both the spectra have been shifted so that their 0-0 transitions are in the origin of the energy axis.

Comparison with experiment

The computed emission spectrum is in good agreement with the experimental one reported by Michl *et al.* (see the fluorescence part of the luminescence spectrum reported figure 1 of reference [136]). The general shape of the spectrum is well reproduced, mainly showing a blue-shift in energy of $\approx$ 1500 cm^{-1}. The main computational error is the underestimation of the 0-0 band intensity, already seen in the calculation of the absorption spectrum by Minaev *et al.* [129], which is due to the small value of the transition dipole moment $\boldsymbol{\mu}_{if}(\mathbf{Q}'_0)$ computed at the TD-DFT level, resulting in a too weak FC contribution to the spectrum.

It should be noted that in our original work [19], a correction of $\boldsymbol{\mu}_{if}$ was performed in order to allow a detailed comparison with the high definition absorption and emission quasi-line bands reported by Gradyushko and coworkers [134, 135]. This scaling procedure is difficult to implement in an automatic fashion and has not been integrated in our procedure.



4.8 Anisole

4.8.1 Introduction

Gas phase spectroscopic studies have been currently focused on the interaction between aromatic molecules, *e.g.* π-π stacking, and on DNA nucleotide base pairing [137–144]. At the same time, a great effort has been devoted to microsolvation studies, dealing with complexes formed by organic molecules and water or other solvents [145–149]. Among aromatic molecules, phenol and anisole are particularly interesting because of the coexistence of prototypical functional groups: the aromatic ring, H-bond donor and acceptor and the methyl group (anisole); yielding to a number of different solvent-molecule interaction schemes (see Figure 4.22): hydrogen bonding, van der Waals forces, dipole-dipole interactions. Moreover, in many cases none of these mechanisms is clearly dominant and a delicate balance of different terms is likely to be expected.

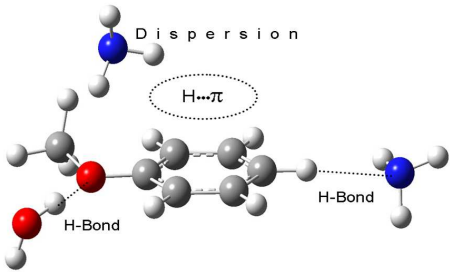


Figure 4.22: Different types of interaction in complexes of anisole.

Phenol represents the prototype of aromatic alcohols and it has been deeply studied as an isolated molecule, as a dimer or as a partner in intermolecular complexes [144, 145, 150]: in all the reported cases the leading interaction has been unambiguously assigned as the formation of a strong hydrogen bond. On the other side, recent studies of the anisole-water [151] and anisole-ammonia [152] adducts highlighted the complexity of the potential energy surface of molecular systems that involve the anisole, showing the different binding properties for these two systems. For anisole-water [146, 147], an hydrogen bond ($OH\cdots O$) is formed with the oxygen of anisole yielding to a planar strcture, while for anisole-ammonia [152,153] interactions with the the delocalized π-electron density of the anisole ring leads to a non planar one.

The vibrationally resolved absorption spectrum of the $S_1 \leftarrow S_0$ electronic transitions of anisole [20] represents an example of the simulation accuracy achievable when good quality geometries and force fields for both electronic states are provided. This can be acheived by applying the Density Functional Theory (DFT) and its Time-Dependent (TD-DFT) counterpart [13, 154] to reliable predictions of geometries and harmonic frequencies of molecular systems in ground and excited electronic state respectively. The electronic transition between the vibrational ground states can be further improved by post-Hartree-Fock single point energy calculations. In particular, methods based on the coupled cluster ansatz are known to provide accurate energies for the ground as well as singly excited electronic states, with computational costs that are reasonable even for medium-size systems. Therefore, an integrated QM approach can take advantage of the overall accuracy of the

energies given by coupled cluster models, while at the same time using the reliable structures and frequencies computed by methods based on density functional theory.

4.8.2 Quantum mechanical calculations

Geometry optimizations of anisole in the ground and first excited electronic state have been carried out using the DFT and TD-DFT [13], respectively. The B3LYP functional has been adopted in conjunction with the 6-311+G(d,p) basis set. While more recent functionals [155, 156] may well perform better than B3LYP, this was deemed sufficiently accurate for the purpose, especially since relative energies are anyway corrected at the coupled cluster level (see below). Harmonic vibrational frequencies have been computed for each structure. For the ground state, the standard procedure based on the availability of analytical second derivatives for DFT was followed; for the excited state, frequencies were evaluated by numerical differentiation of the analytical TD-DFT energy gradients. Anharmonic perturbative [103] calculations have been performed at the B3LYP/6-311+G(d,p) level for the ground electronic state, and subsequently used to account for anharmonicity in both S_0 and S_1 electronic states, as described in section 3.3.1. The zero point vibrational energy (ZPE) can be expressed [103] as

$$ZPE = \frac{1}{2}(ZPE_{harm} + ZPE_{anh}) + \xi_0 - \frac{1}{4}\sum_i \xi_{ii}, \tag{4.3}$$

where $ZPE_{harm} = \frac{1}{2}\sum_i \omega_i$ and $ZPE_{anh} = \frac{1}{2}\sum_i \vartheta_i$, $\boldsymbol{\omega}$ and $\boldsymbol{\vartheta}$ being the harmonic and anharmonic fundamental vibrational frequencies, respectively (see Ref. [103] for the definition of ξ_0 and ξ_{ii}). For the ground electronic state, the anharmonic zero point energy correction has been computed according to equation 4.3, as implemented in the GAUSSIAN package [18,103]. For the ground state the last two terms of equation 4.3 account for just about 0.2 % ($\approx$ 70 cm^{-1}) of the total ZPE. Thus for the excited state the ZPE has been approximated as an average between ZPE$_{harm}$ and ZPE$_{anh}$. All calculations rooted in the density functional theory have been performed with the GAUSSIAN suite of quantum chemistry programs [18].

The energy separation between ground and excited electronic state has been refined by single-point energy calculations at the coupled cluster level of theory including single and double excitations (CCSD) [113, 114], with the 6-311+G(d,p) and the aug-cc-pVDZ (AVDZ) [111, 112] basis sets. For the S_1 first excited electronic state the equation-of-motion CCSD counterpart (EOM-CCSD) [115] has been exploited. The CCSD and EOM-CCSD calculations have been performed with the MOLPRO [116] package.

4.8.3 Results and Discussion

Structure of anisole in S_0 and S_1 states

The planarity of anisole in both the ground and the first excited electronic states has been determined by the high resolution electronic excitation spectrum of the band origin in the $S_1 \leftarrow S_0$ electronic transition using laser-induced fluorescence (LIF) spectroscopy [157]. Table 4.7 compares theoretical and experimental rotational constants, showing an average deviation of about 0.5 % for both electronic states.

	calc	exp [157]
	S_0	
A	0.168591	0.167745
B	0.052032	0.052349
C	0.040065	0.040222
	S_1	
A	0.161597	0.159978
B	0.051761	0.051896
C	0.039500	0.039515

Table 4.7: Comparison of calculated and experimental rotational constants (in cm^{-1}) for anisole in its ground and excited electronic state.

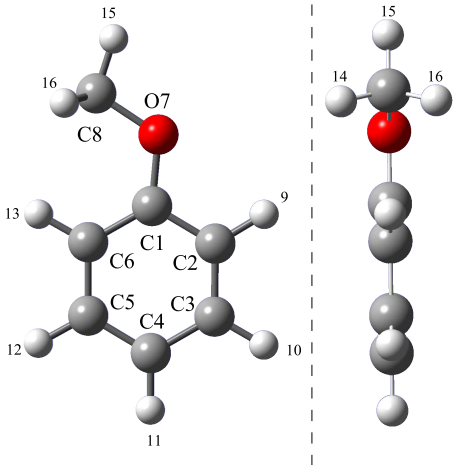


Figure 4.23: Geometry and numbering scheme of anisole.

The agreement is remarkable, and confirms the good quality of the calculated geometries. For the ground state, accurate structures calculated using the density functional theory with the B3LYP functional and basis set of triple$-\zeta$ quality have already been reported [158, 159], but for the excited state the TD-B3LYP results from this work should be considered as the best estimates available at present. The structure and atom numbering of anisole are shown in Figure 4.23, and the geometric parameters for the ground and first excited electronic states are listed in Table 4.8. The $S_1 \leftarrow S_0$ electronic transition has a mixed n $\rightarrow \pi^*$/ $\pi \rightarrow \pi^*$ character. The frontier Kohn-Sham orbitals (MO) involved in the n $\rightarrow \pi^*$ / $\pi \rightarrow \pi^*$ transition, computed at the B3LYP/6-311+G(d,p) level for the S_1 geometry, are shown in Figure 4.32, and the electron density difference between the ground and excited states is sketched in Figure 4.24. TD-DFT calculations show that the first electronic excitation is dominated ($\geq$60%) by the HOMO-LUMO transition, with some contributions from the HOMO-1 to the LUMO+1 transitions. On inspection of the MO plots, it clearly appears that the most important effect related to the HOMO-LUMO transition is the transfer of electron density (ELD) from the oxygen atom to the aromatic ring, and this has been also confirmed by density-difference plots.

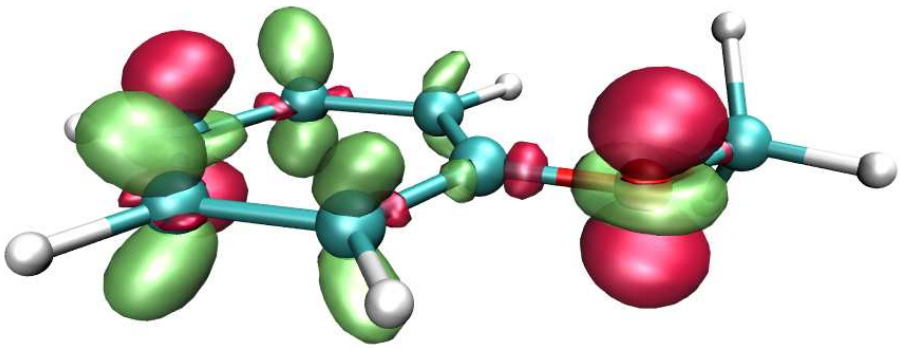


Figure 4.24: The S_1-S_0 difference density for the anisole. The electronic density corresponds to the ground and first excited states computed for the structure optimized at TD-B3LYP/6-311+G(d,p) level. The regions that have lost electron density as a result of the transition are shown in magenta, and the regions with the lime color gained electron density. The isosurfaces shown here correspond to the frontier 0.0035.

The transfer of electron density from the oxygen atom to the aromatic ring leads to some changes in the structure, but does not affect the molecular symmetry. On the whole, all CC bonds become longer in the excited state. As for the CO bonds, C1O7 becomes shorter by 0.017 Å and O7C8 longer by 0.009 Å in the excited state. Other structural changes are related to the increase of CCC angles: C3C4C5 and C6C1C2 by ≈ 3.5 deg, and C1O7C8 by 2.4 deg. These changes are in line with a decrease of electron density on the oxygen atom, and an increase on the aromatic ring. Earlier calculations at the configuration interaction singles level (CIS/6-311G(d,p)) [159] predicted both CO bonds about 0.02 Å shorter with respect to TD-B3LYP, and gave the opposite trend for the change of the O7C8 bond (see supplementary Table 4.14). On the whole, the geometry of anisole does not change much during the transition, as required for applicability of the Franck-Condon approximation.

	S_0	S_1	$\Delta(S_1 - S_0)$
C1C2	1.401	1.424	0.024
C2C3	1.388	1.425	0.037
C3C4	1.398	1.412	0.015
C4C5	1.390	1.418	0.028
C5C6	1.398	1.426	0.028
C6C1	1.397	1.421	0.024
C1O7	1.366	1.349	-0.017
O7C8	1.420	1.429	0.009
C2H9	1.083	1.081	-0.003

Table 4.8: Selected geometrical parameters (bond lengths in Å and angles in degrees) of anisole in its ground and excited state, computed at B3LYP/6-311+G(d,p) and TD-B3LYP/6-311+G(d,p) level.

Table 4.8: Continuation

	S_0	S_1	$\Delta(S_1 - S_0)$
C3H10	1.084	1.081	-0.003
C4H11	1.083	1.085	0.001
C5H12	1.084	1.081	-0.003
C6H13	1.082	1.080	-0.002
C8H14(16)	1.096	1.093	-0.003
C8H15	1.089	1.088	-0.001
C1O7C8	118.6	121.0	2.4
C2C1O7	115.7	113.9	-1.8
C6C1O7	124.5	122.9	-1.6
C1C2C3	120.0	118.1	-1.9
C2C3C4	120.6	119.1	-1.5
C3C4C5	119.2	122.4	3.3
C4C5C6	120.9	119.4	-1.5
C5C6C1	119.5	117.7	-1.7
C6C1C2	119.8	123.2	3.4
H15C8O7	105.9	105.5	-0.3
H14C8O7	111.4	110.6	-0.9
H15C8O7C1	180.0	180.0	0.0
H14C8O7C2	61.3	61.0	-0.2
H16C8O7C2	-61.3	-61.0	0.2

Table 4.8: Selected geometrical parameters (bond lengths in Å and angles in degrees) of anisole in its ground and excited state, computed at B3LYP/6-311+G(d,p) and TD-B3LYP/6-311+G(d,p) level.

Vibrational frequencies in the ground electronic state of anisole

Preliminary to analysis of the vibrational structure of the $S_1 \leftarrow S_0$ transition, the computed vibrational frequencies of the ground state must be properly assigned. In line with previous works [158, 160], we chose to follow the labeling convention proposed by Balfour [161] to refer to the vibrational modes. For the aromatic ring modes, the Wilson notation [162] is used throughout the paper. Cardinal numbers are used to number the 42 normal modes of anisole, while the notation v_n refers to the assignment. Experimental and calculated frequencies of anisole in its ground electronic state are listed in Table 4.9 along with the proposed assignment. In discussing the assignments, we will follow the scheme given by Varsanyi [163].

mode	sym	exp	calc			assignment
		[161]	[158, 160][a)]	[159][b)]	This work[c)]	
1	A”	81.5	90	92	86	$v'_{\text{COC torsion}}$
2	A”	209	203	203	200	v'_{10b}
3*	A’	260	250	251	253	$v'_{\text{COC bending}}$
4	A”	263	266	267	258	$v'_{\text{O-CH}_3\text{ torsion}}$
5	A”	415	412	416	417	v'_{16a}
6*	A’	430[d)]	433	437	438	v'_{9b}
7	A”	511	502	509	511	v'_{16b}
8	A’	553	543	549	554	v'_{6a}
9	A’	618	610	616	624	v'_{6b}
10	A”	690	669	689	697[e)]	v'_4
11	A”	752	738	750	757	v'_{11}
12	A’	788	777	782	784	v'_1
13	A”	819	806	811	812	v'_{10a}
14	A”	880	867	877	883	v'_{17b}
15	A”	956	941	946	959	v'_{17a}
16	A”	975	956	965	982	v'_5
17	A’	997	981	988	995	v'_{12}
18	A’	1022	1015	1020	1019	v'_{18a}
19	A’	1039	1041	1045	1040	$v'_{\text{O-CH}_3\text{ stretch.}}$
20*	A’	1073	1075	1079	1083	v'_{18b}
21	A”	1143	1138	1146	1147	$v'_{\text{CH}_3\text{ rock.}}$
22	A’	1151	1147	1153	1164	v'_{15}
23	A’	1169	1165	1169	1178	v'_{9a}
24	A’	1180	1173	1179	1177	$v'_{\text{CH}_3\text{ rock.}}$
25*	A’	1253	1248	1249	1239	v'_{7b}
26	A’	1292	1305	1306	1306	v'_3
27	A’	1332	1330	1328	1336	v'_{14}
28	A’	1442	1437	1445	1468	$v'_{\text{CH}_3\text{ sym. def.}}$
29	A’	1455	1449	1456	1455	v'_{19b}
30	A”	1452	1469	1475	1480	$v'_{\text{CH}_3\text{ antisym. def.}}$
31	A’	1464	1456	1460	1466	$v'_{\text{CH}_3\text{ antisym. def.}}$
32	A’	1497	1492	1498	1495	v'_{19a}

Table 4.9: Calculated and experimental frequencies (in cm^{-1}) of anisole in its ground electronic state. Modes for which a revised assignment is proposed are marked by an asterisk.

[a]B3LYP/6-311++G(d,p) harmonic frequency calculations, frequencies up to 2000 cm^{-1} scaled by a factor of 0.973 and those of higher frequency range by the factor 0.963.

[b]B3LYP/6-311G(d,p) harmonic frequency calculations, scaled by a factor of 0.979.

[c]B3LYP/6-311+G(d,p) anharmonic frequency calculations.

[d]Ref. [159].

[e]Harmonic frequency value.

Table 4.9: Continuation

mode	sym	exp	calc			assignment
		[161]	[158, 160][a)]	[159][b)]	This work[c)]	
33	A'	1588	1542	1592	1583	v'_{8b}
34	A'	1599	1606	1612	1601	v'_{8a}
35	A'	2900	2903	2934	2870	$v'_{CH_3 \text{ sym. stretch.}}$
36	A"	2942	2964	2990	2903	$v'_{CH_3 \text{ asym. stretch.}}$
37	A'	3004	3034	3063	2994	$v'_{CH_3 \text{ asym. stretch.}}$
38*	A'	3026	3063	3101	3007	v'_{13}
39*	A'	3037	3070	3094	3042	v'_{7a}
40*	A'	3062	3089	3135	3047	v'_{20a}
41*	A'	3092	3093	3125	3069	v'_{20b}
42*	A'	3105	3101	3117	3078	v'_{2}

Table 4.9: Calculated and experimental frequencies (in cm^{-1}) of anisole in its ground electronic state. Modes for which a revised assignment is proposed are marked by an asterisk.

Among the tangential in-plane vibrations, Balfour [161] assigned the 1151 cm^{-1}, 1073 cm^{-1} and 260 cm^{-1} bands to modes v'_{9b}, v'_{15} and v'_{18b}, respectively. Hoffman et al. [160] proposed to modify the original assignment into modes v'_{15}, v'_{9b} and v'_{18b}. Inspection of the displacements underlying the normal modes computed here suggests still another correction, namely v'_{15}, v'_{18b} and $v'_{\text{COC bending}}$, respectively. We also propose to assign the vibration at 430 cm^{-1}, not observed by Balfour [161], but reported recently by Matsumoto et al. [159], to mode v'_{9b}. Concerning the radial in-plane modes, some ambiguities remain about the assignment of the *C-X(H) stretching* modes (v'_2, v'_{7a}, v'_{7b}, v'_{20a}, v'_{20b}, v'_{13}), which are related to the band at 1253 cm^{-1} and to five high frequency bands (over 3000 cm^{-1}). The absorption at 1253 cm^{-1} has been first assigned by Balfour to the v'_{7a} mode, but Matsumoto et al. changed it to the mode v'_{13}, while Hoffman et al. supported the original Balfour's assignment. We propose here to reassign this mode to v'_{7b}; this entails also a reassignment of the *C-H stretching* modes, as listed in Table 4.9. As for the out-of-plane vibrations, all earlier assignments are confirmed. Finally, concerning the vibrations of the methoxy group, we suggest to assign the 260 cm^{-1} frequency to the *COC bending* mode, as discussed above.

In the present work anharmonic frequency calculations for anisole in its ground state have been performed for the first time. The results show a very good agreement with the experimental data. In particular, a RMS deviation of 8 cm^{-1} between computed and experimental frequencies is achieved, provided frequencies related to the highly anharmonic modes (involving C-H and CH_3 vibrations) are excluded. Nevertheless, even for strongly anharmonic modes the maximum deviation does not exceed 40 cm^{-1}, and the RMS value computed with inclusion of all modes is still only 13 cm^{-1}. Comparable overall agreements have only been achieved by adopting different scaling factors for the high- and low-frequency portions of the spectrum [160]. However, in this case the relative error

[a] B3LYP/6-311++G(d,p) harmonic frequency calculations, frequencies up to 2000 cm^{-1} scaled by a factor of 0.973 and those of higher frequency range by the factor 0.963.

[b] B3LYP/6-311G(d,p) harmonic frequency calculations, scaled by a factor of 0.979.

[c] B3LYP/6-311+G(d,p) anharmonic frequency calculations.

affecting a particular frequency is not easily related to its nature, in the sense that an anharmonic mode like v_2' can show a small deviation, while at the same time a ring deformation like v_{8b}' is affected by a large error (46 cm^{-1}). Overall, it would appear that the anharmonic frequency calculations for anisole presented here are the most reliable theoretical values reported to date: small residual discrepancies, where present, are clearly related to the physical nature of the phenomenon and are not accidental. From another viewpoint, this also provides an indirect support to the use of theoretical anharmonic corrections for excited state calculations.

Harmonic and anharmonic frequencies in the S_1 excited electronic state of anisole

To compare simulated spectra with experimental data, one needs to account for the anharmonicity of vibrational modes in the first singlet excited state. Due to the lack of analytical second derivatives, direct computation of anharmonic frequencies of anisole at the TD-DFT level is not practical. Here we introduce two approaches to obtain anharmonic correction in the S_1 state. Both are based on the Duschinsky transformation [31] used to represent the excited state normal modes, and make use of ground state anharmonic frequencies. The calculated ($\boldsymbol{\alpha}_{\mathrm{TA}}$) or empirical ($\boldsymbol{\alpha}_{\mathrm{EA}}$) mode-specific scaling factors are derived (see section 3.3.1) from comparison of the calculated harmonic frequencies to anharmonic calculations or to experimental data, respectively. In view of the good agreement between theoretical and experimental ground state frequencies that has been highlighted before, in the present instance the two sets of values (Table 4.10) are actually close to each other, with the exception of a few strongly anharmonic modes.

	S_0				S_1		
Mode	ω'	ϑ'	α_{TA}	α_{EA}	ω''	$\vartheta''_{\mathrm{TA}}$	$\vartheta''_{\mathrm{EA}}$
1	90	86	0.956	0.902	72	69	66
2	204	200	0.982	1.024	86	85	85
3	256	253	0.987	1.016	139	136	140
4	268	258	0.963	0.983	199	192	197
5	422	417	0.989	0.984	252	249	256
6	446	438	0.981	0.964	376	376	372
7	517	511	0.988	0.988	437	428	421
8	560	554	0.989	0.987	441	433	436
9	629	624	0.992	0.982	517	512	509
10	697	697[f)]	1.000	0.990	530	530	524
11	764	757	0.991	0.985	546	540	538
12	797	784	0.984	0.988	604	596	595
13	827	812	0.981	0.990	629	623	620
14	894	883	0.988	0.985	680	673	670
15	971	959	0.987	0.984	781	769	772

Table 4.10: Calculated harmonic and anharmonic frequencies (in cm^{-1}) and mode specific scaling factors.

[f]Harmonic frequency value.

Table 4.10: Continuation

		S_0			S_1		
Mode	ω'	ϑ'	α_{TA}	α_{EA}	ω''	ϑ''_{TA}	ϑ''_{EA}
16	989	982	0.992	0.985	858	850	845
17	1009	995	0.985	0.988	973	957	960
18	1041	1019	0.979	0.982	990	970	972
19	1065	1040	0.977	0.976	1006	986	982
20	1102	1083	0.982	0.973	1035	1014	1008
21	1169	1147	0.981	0.978	1143	1121	1118
22	1178	1164	0.988	0.977	1152	1136	1127
23	1194	1178	0.986	0.979	1161	1144	1136
24	1201	1177	0.980	0.982	1183	1161	1161
25	1271	1239	0.974	0.986	1278	1247	1258
26	1335	1306	0.978	0.968	1310	1284	1277
27	1357	1336	0.985	0.982	1395	1369	1365
28	1474	1468	0.996	0.978	1420	1390	1389
29	1485	1455	0.980	0.980	1449	1434	1417
30	1493	1480	0.991	0.972	1457	1423	1424
31	1506	1466	0.974	0.972	1476	1450	1446
32	1527	1495	0.979	0.981	1483	1470	1443
33	1624	1583	0.974	0.978	1494	1458	1454
34	1642	1601	0.975	0.974	1521	1483	1482
35	3003	2870	0.956	0.966	3032	2897	2928
36	3060	2903	0.948	0.961	3101	2940	2981
37	3132	2994	0.956	0.959	3153	3014	3025
38	3163	3007	0.951	0.957	3171	3024	3042
39	3170	3042	0.960	0.958	3201	3072	3066
40	3187	3047	0.956	0.961	3208	3061	3078
41	3194	3069	0.961	0.968	3223	3095	3117
42	3204	3078	0.961	0.969	3229	3101	3128

Table 4.10: Calculated harmonic and anharmonic frequencies (in cm^{-1}) and mode specific scaling factors.

It should be pointed out that TD-DFT frequencies differ significantly from their CIS counterparts (see supplementary Table 4.15). In particular, discrepancies of over 50 cm^{-1} are found for modes 2, 8, 33, 34, 41 and 42 of anisole. Since previous assignments of the fundamental vibrational transitions in S_1 have been based on CIS calculations [159,160], one can expect that some reassignments are needed. As discussed above, the electronic excitation does not affect significantly the molecular structure; however, it leads to important changes in vibrational properties. These are related to changes of normal modes (Duschinsky rotation), as well as to significant frequency variations for some modes. Table 4.11 lists the assignments and the calculated anharmonic frequencies for both electronic states.

S_0		S_1		Δ S_1-S_0
ϑ' [g]	assignment	ϑ''_{EA} [h]	assignment	
86	$v'_{COC\ torsion}$	66	0.89 $v'_{COC\ torsion}$	-20
417	v'_{16a}	85	-0.83 v'_{16a}	-332
200	v'_{10b}	140	0.75 v'_{10b} [+ 0.17 $v'_{O\text{-}CH_3\ torsion}$]	-60
258	$v'_{O\text{-}CH_3\ torsion}$	197	0.74 $v'_{O\text{-}CH_3\ torsion}$ [- 0.22 v'_{10b}]	-61
253	$v'_{COC\ bending}$	256	$v'_{COC\ bending}$	3
511	v'_{16b}	372	-0.49 v'_{16b} + 0.46 v'_4	-139
438	v'_{9b}	421	0.97 v'_{9b}	-17
812	v'_{10a}	436	0.75 v'_{10a} [- 0.16 v'_{17a}]	-376
624	v'_{6b}	509	-0.61 v'_{6b} - 0.39 v'_{6a}	-115
697 [i]	v'_4	524	0.48 v'_{16b} + 0.48 v'_4	-173
554	v'_{6a}	538	-0.60 v'_{6a} + 0.38 v'_{6b}	-16
959	v'_{17a}	595	0.74 v'_{17a}	-364
757	v'_{11}	620	-0.77 v'_{11}	-137
883	v'_{17b}	670	0.66 v'_{17b} [- 0.29 v'_5]	-213
784	v'_1	772	v'_1	-12
982	v'_5	845	0.57 v'_5 [+ 0.27 v'_{17b}]	-136
995	v'_{12}	960	-0.70 v'_{12} [+ 0.30 v'_{18a}]	-35
1019	v'_{18a}	972	-0.58 v'_{18a} [- 0.22 v'_{12} - 0.17 $v'_{O\text{-}CH_3\ stretch.}$]	-47
1040	$v'_{O\text{-}CH_3\ stretch.}$	982	-0.43 v'_{18b} + 0.39 $v'_{O\text{-}CH_3\ stretch.}$	-58
1083	v'_{18b}	1008	0.52 v'_{18b} + 0.44 $v'_{O\text{-}CH_3\ stretch.}$	-74
1147	$v'_{CH_3\ rock.}$	1118	$v'_{CH_3\ rock.}$	-29
1164	v'_{15}	1127	0.55 v'_{15} + 0.30 v'_{9a}	-37
1178	v'_{9a}	1136	0.67 v'_{9a} [- 0.21 v'_{15}]	-42
1177	$v'_{CH_3\ rock.}$	1161	0.80 $v'_{CH_3\ rock.}$ [- 0.17 v'_{15}]	-16
1239	v'_{7b}	1258	0.86 v'_{7b}	19
1306	v'_3	1277	0.56 v'_3 + 0.30 v'_{14}	-29
1336	v'_{14}	1365	-0.50 v'_{14} [+ 0.18 v'_3 + 0.16 v'_{8a}]	29
1455	v'_{19b}	1389	0.75 v'_{19b}	-65
1468	$v'_{CH_3\ sym.\ def.}$	1417	0.66 $v'_{CH_3\ sym.\ def.}$ [+ 0.21 v'_{19a}]	-49
1601	v'_{8a}	1424	-0.49 v'_{8a} [+ ...]	-177
1495	v'_{19a}	1446	-0.64 v'_{19a} [+ 0.21 $v'_{CH_3\ sym.\ def.}$]	-49
1466	$v'_{CH_3\ antisym.\ def.}$	1443	$v'_{CH_3\ antisym.\ def.}$	-23
1480	$v'_{CH_3\ antisym.\ def.}$	1454	0.87 $v'_{CH_3\ antisym.\ def.}$	-26
1583	v'_{8b}	1482	-0.81 v'_{8b} [+ 0.16 v'_{8a}]	-101

Table 4.11: Assignments and calculated frequencies (in cm^{-1}) for anisole in its ground and first excited electronic state.

[g] S_0 frequencies are listed in such sequence as to to match the corresponding S_1 frequencies.
[h] S_1 frequencies are listed in order of increasing harmonic frequency, see ω' in Table 4.10.
[i] Harmonic frequency value.

Table 4.11: Continuation

S_0		S_1		Δ S_1-S_0
$\vartheta'^{g)}$	assignment	$\vartheta''_{EA}{}^{h)}$	assignment	
2870	$v'_{CH_3 \text{ sym. stretch.}}$	2928	$v'_{CH_3 \text{ sym. stretch.}}$	58
2903	$v'_{CH_3 \text{ asym. stretch.}}$	2981	$v'_{CH_3 \text{ asym. stretch.}}$	78
2994	$v'_{CH_3 \text{ asym. stretch.}}$	3025	$v'_{CH_3 \text{ asym. stretch.}}$	31
3007	v'_{13}	3042	$0.54\ v'_{13} + 0.36\ v'_{20a}$	35
3042	v'_{7a}	3066	v'_{7a}	24
3047	v'_{20a}	3078	$-0.49\ v'_{20a} + 0.44\ v'_{13}$	31
3069	v'_{20b}	3117	$0.80\ v'_{20b}$	48
3078	v'_{2}	3128	$0.92\ v'_{2}$	50

Table 4.11: Assignments and calculated frequencies (in cm^{-1}) for anisole in its ground and first excited electronic state.

In general, normal modes in the ground and excited electronic states are not identical. In particular modes 6, 9-11, 19-20, 38 and 40 in S_1 are described by approximately equal contributions from pairs of S_0 vibrations. Even for the modes less affected by Duschinsky rotation, significant changes in vibrational energy may occur: thus, v'_{16a}, v'_{10a}, v'_{17a}, v'_{11} and v'_{8b} are shifted by more than 100 cm^{-1}.

Simulated vs experimental spectra

The simulated vibronic profile of the $S_1 \leftarrow S_0$ electronic transition is compared to the experimental data from Resonance Enhanced Multi-Photon Ionization (REMPI) spectra [160]. In Figure 4.25, the best fully theoretical spectrum is shown on an absolute energy scale.

This has been calculated by the Franck-Condon Herzberg-Teller approach, with the frequencies corrected for anharmonicity, and the energy of the electronic transition adjusted based on (CCSD/EOM-CCSD)/aug-cc-pVDZ single point calculations. In this way a very good agreement (0.05 eV) between theoretical and experimental results for the 0-0 electronic transition has been achieved. The DFT/TD-DFT route also yields results close (0.2 eV) to the experiment (see Table 4.7).

	calc	exp [157]
TD-B3LYP/6-311+G(d,p)	4.694	4.511
EOM-CCSD/6-311+G(d,p)	4.642	
EOM-CCSD/AVDZ	4.564	

Table 4.12: Comparison of calculated and experimental energy of the $S_1 \leftarrow S_0$ transition (in eV) for anisole.

Nevertheless, such an accuracy is still insufficient to discuss the vibrational states in S_1, and it is

[g] S_0 frequencies are listed in such sequence as to to match the corresponding S_1 frequencies.
[h] S_1 frequencies are listed in order of increasing harmonic frequency, see ω' in Table 4.10.

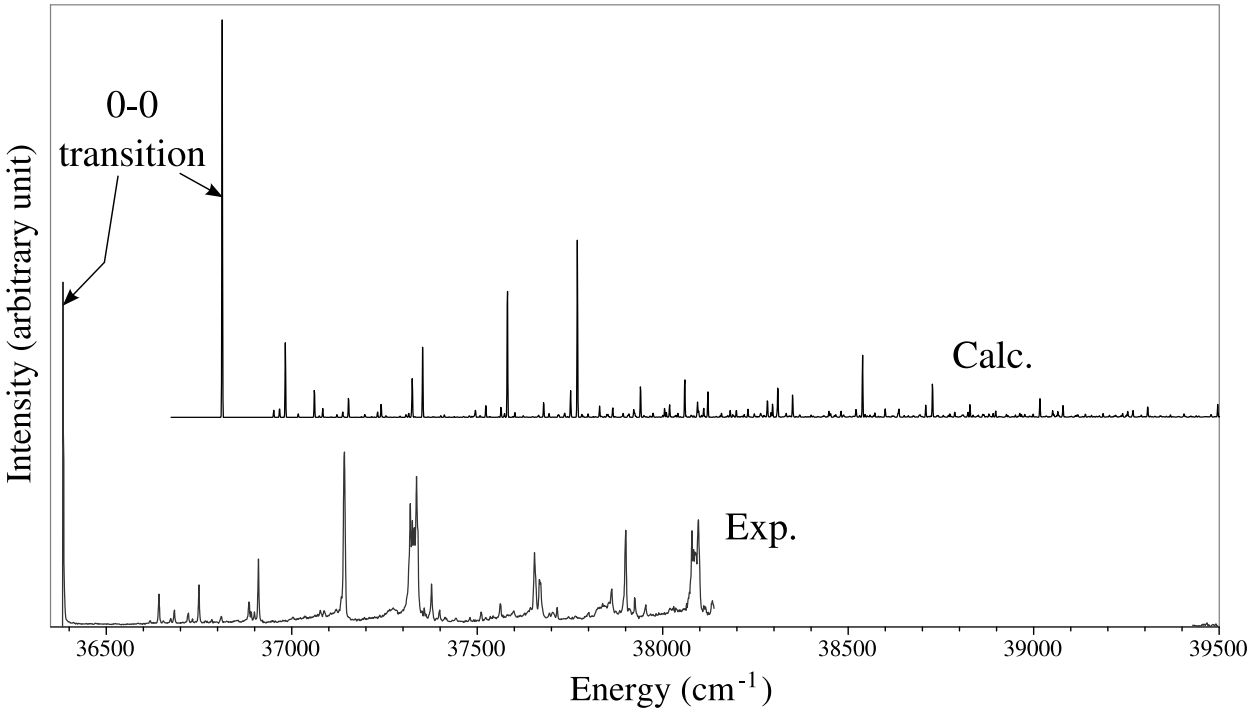


Figure 4.25: Theoretical and experimental [160] spectra of the $S_1 \leftarrow S_0$ transition of anisole as a function of absolute energy (in cm^{-1}).

therefore necessary to compare the values relative to the 0-0 origin: such a comparison is presented in Figure 4.26.

Both simulated spectra are calculated by the Franck-Condon approximation and include the anharmonicity corrections, but differ in the (arbitrary) values adopted for the Full-Width at Half-Maximum (FWHM) (2 and 20 cm^{-1}, respectively). Since the REMPI spectra of anisole show sharp bands, the best match is achieved with a FWHM of 2 cm^{-1}, and this value has been applied to produce all other spectra shown in the following.

Figure 4.33, situated at the end of this section, shows how the approximations implemented to model theoretical spectra affect the calculated vibrational profiles, and more in general the accuracy of the results.

Panel A displays the spectrum in a range 0-1800 cm^{-1}, which correspond to the full range of the experimental spectra, while an expansion of the 800-1800 cm^{-1} region is presented in panel B. At the first level of approximation, the Duschinsky rotation between normal modes, as well as any frequency change between the states, are neglected [$\mathbf{J} = \mathbf{I}; \omega$'(i)=$\omega$(i)]. In other words, this approach takes into account only the translation between the ground and exited state geometries and normal modes ($\mathbf{K} \neq 0$), and thus corresponds roughly to the LCM approximation. It should be noted that LCM [11] estimates displacements from the PES gradient at the ground-state equilibrium geometry (accounting in principle for anharmonicity to the first order) [11, 96, 164, 165], while our approach retain the exact displacements from the optimized ground- and excited-state geometries. Next the Franck-Condon spectra obtained with the the Duschinsky rotation are reported, and finally the effect of the anharmonic correction is shown. The spectrum corresponding to the Franck-Condon Herzberg-Teller approximation is not plotted separately, since on the adopted scale it would be almost superimposable to its Franck-Condon only counterpart. From Figure 4.33 it is evident that

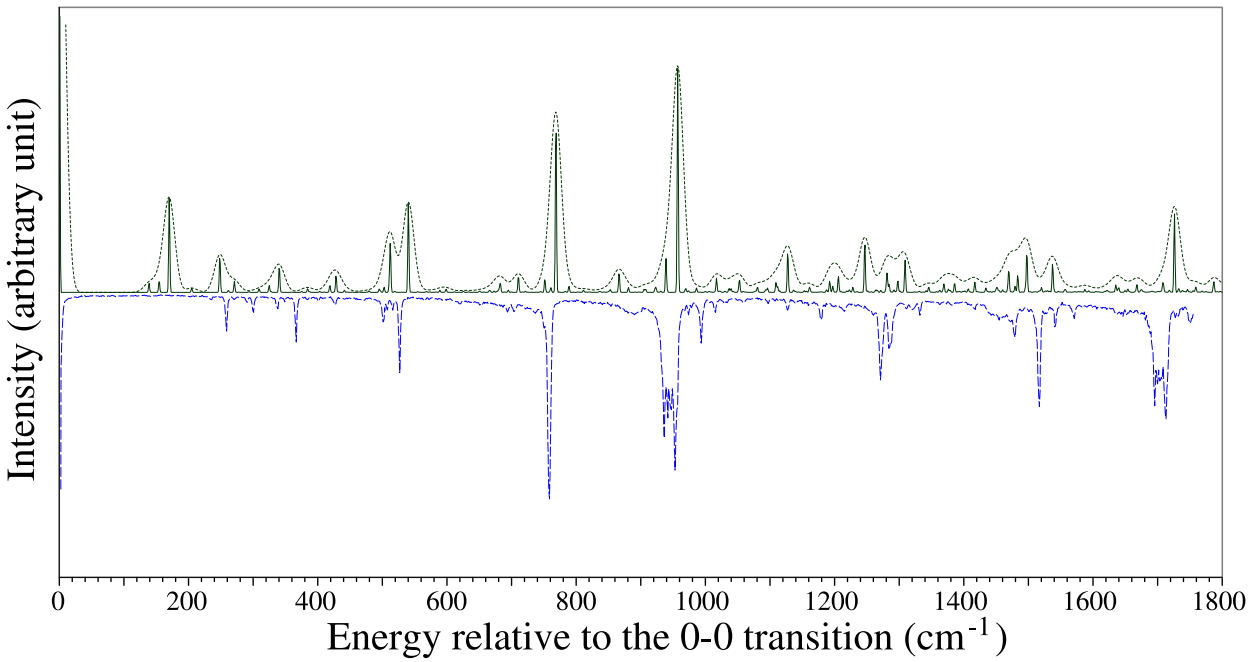


Figure 4.26: Theoretical (upper panel) and experimental [160] (lower panel) spectra of the $S_1 \leftarrow S_0$ transition of anisole. Solid line, spectrum calculated with a FWHM of 2 cm^{-1}; dashed line, spectrum calculated with a FWHM of 20 cm^{-1}.

the LCM approach provides only a rough estimation of the actual spectra. The two most intense bands of the REMPI spectrum are well reproduced, but the third band at $\approx$ 550 cm^{-1} shows similar intensity, at variance with the REMPI data. The differences are even more pronounced for the 800-1800 cm^{-1} spectral range, where the LCM approximation misses an intense band at $\approx$ 1700 cm^{-1}, as well as many other weak features. The only spectral region where the LCM shows a better agreement to the experimental spectra is the (essentially featureless) 0-200 cm^{-1} range: here, the spectrum based on TD-DFT geometry and frequencies for the excited state shows an intense band at $\approx$ 170 cm^{-1}. The band corresponds to the first vibrational overtone of the 2nd vibration of S_1, and therefore the discrepancy can be traced back to an inappropriate description of the strongly anharmonic out-of-plane ring deformation v'_{16a} coupled with the *O-CH$_3$ torsion*. However, in the 200-1800 cm^{-1} range the Franck-Condon calculations provide an accurate intensity pattern for all the important bands of the REMPI spectra. Introduction of the correction for anharmonicity causes a red-shift of the frequencies, and results in a good overall agreement for the band positions as well. It is worthwhile underlining that the Franck-Condon calculations from our work are able to reproduce the rich vibrational structure of the REMPI spectra. As a matter of fact, since the $S_1 \leftarrow S_0$ electronic transition of anisole is dipole-allowed ($\mu(\mathbf{Q}'_0) \approx 0.6$ a.u.), the Franck-Condon approximation is sufficient to reproduce all the main spectral features, and the more computationally demanding Franck-Condon Herzberg-Teller calculations do not introduce significant improvements.

Vibrational bands in the S_1 excited electronic states of anisole

There have been several attempts to assign fundamental vibrations of anisole in its first electronic excited state based on the vibronic structure of the $S_1 \leftarrow S_0$ electronic transition, starting from early

work by Balfour [166], to the recent papers by Matsumoto et al. [159] and Hoffman et al. [160]. In these latter, computations of excited state frequencies at the CIS level of theory have also been used to assist the assignment. Nevertheless, some of the assignments must still be regarded as tentative, the main difficulties being related to low signal intensity, as well as to the high density of possible fundamental bands, not to mention overtones and binary combinations. In principle, simulations of spectra give insight into the vibrational components of each of the vibronic transitions, and therefore provide a straightforward route to assignment.

In the present work a direct comparison between experimental and simulated spectra has been used as the basis for the assignment of the vibrational modes in the excited state. It is immediately apparent from Figure 4.26 that the main features of the experimental REMPI spectrum are well reproduced by the simulation. For example, one of the most intense bands, at 759 cm^{-1}, has been assigned to mode v'_1. The simulated spectrum displays a strong band at 769 cm^{-1} (772 cm^{-1} if the empirical anharmonic correction is adopted), which corresponds to the 15th mode of S_1, indeed assigned as v'_1. The reliability of the theoretical approach can also be gauged by a comparison between the experimental and computed frequencies of the set of the six most intense bands (highlighted in Table 4.13), which results in an RMSD of 15 cm^{-1}.

exp[j]	ϑ''_{TA}	ϑ''_{EA}	assignment[k]
234[l]	204	205	$\{0.89\ v'_{\text{COC torsion}}\}+\{0.75\ v'_{10b}+ 0.17\ v'_{\text{O-CH}_3\text{ torsion}}\}$
259[m]	249	256	$\{v'_{\text{COC bending}}\}$
367	384	393	$\{0.74\ v'_{\text{O-CH}_3\text{ torsion}}- 0.22\ v'_{10b}\}^2$
427	428	421	$\{0.97\ v'_{9b}\}$
501[m]	512	509	$\{-0.61\ v'_{6b} - 0.39\ v'_{6a}\}$
508			combination
516			combination
527[m]	540	538	$\{-0.60\ v'_{6a} + 0.38\ v'_{6b}\}$
621			combination
667			combination
704			combination
750	752	744	$\{-0.49\ v'_{16b} + 0.46\ v'_4\}^{2\ n)}$
759[m]	769	772	$\{v'_1\}$
937[m,o]	922	922	$\{v'_{17a}\}+\{\dots\}$

Table 4.13: Assignments of $S_1 \leftarrow S_0$ vibronic transitions of anisole. All energies are relative to the 0-0 origin. Frequencies are in cm^{-1}. Experimental frequencies that have been extracted from the plots of REMPI spectra reported in Ref. [160] are marked by asterisk.

[j] Experimental data from reference [160].
[k] In parenthesis the fundamental modes of S_1 resulting from the Duschinsky rotation between S_0 vibrations.
[l] Experimental data from reference [159].
[m] Bands considered for the estimation of the RMS deviation.
[n] Alternative assignment: $\{-0.83\ v'_{16a}\}+\{-0.61\ v'_{6b} - 0.39\ v'_{6a}\}$ combination.
[o] Average of four bands considered for the estimation of the RMS deviation.

Table 4.13: Continuation

exp[j]	$\vartheta''_{\mathrm{TA}}$	$\vartheta''_{\mathrm{EA}}$	assignment[k]
943[m,o]	938	941	$\{v'_{17a}\}+\{\dots\}$
948[m,o]			combination
954[m,o]	957	960	$\{\text{-}0.70\ v'_{12}\}$
994	986	982	$\{\text{-}0.43\ v'_{18b} + 0.39\ v'_{\text{O-CH}_3\text{ stretch.}}\}$
1016	1016	1027	$\{v'_{17a}\}+\{\dots\}$
1126	1127	1128	$\{v'_{12}\}+\{\dots\}$
$\approx$ 1140*	1136	1127	$\{0.55\ v'_{15} + 0.30\ v'_{9a}\}$[p]
1152	1160	1161	$\{0.80\ v'_{\text{CH}_3\text{ rock.}}\}$
1179	1191	1189	$\{0.74\ v'_{17a}\}^2$
1271	1247	1258	$\{0.86\ v'_{7b}\}$
1288	1284	1277	$\{0.56\ v'_{3} + 0.30\ v'_{14}\}$
1415			combination
1443	1433	1417	$\{0.66\ v'_{\text{CH}_3\text{ sym. def.}} + 0.21\ v'_{19a}\}$
1455			combination
1479	1483	1482	$\{\text{-}0.81\ v'_{8b}\}$[q]
1517	1497	1498	$\{v'_{12}\} + \{\text{-}0.60\ v'_{6a} + 0.38\ v'_{6b}\}$
$\approx$ 1540*	1536	1544	$\{v'_{1}\}^2$
1571			combination
1636			combination
1713[m,r]	1726	1732	$\{v'_{1}\} + \{v'_{12}\}$
2967	2897	2928	$\{v'_{\text{CH}_3\text{ sym. stretch.}}\}$
2979			combination
2990			combination
3049	3024	3042	$\{0.54\ v'_{13} + 0.36\ v'_{20a}\}$
3076	3061	3078	$\{\text{-}0.49\ v'_{20a} + 0.44\ v'_{13}\}$
3084			combination
3099	3095	3117	$\{0.80\ v'_{20b}\}$
3107	3101	3128	$\{0.92\ v'_{2}\}$

Table 4.13: Assignments of $S_1 \leftarrow S_0$ vibronic transitions of anisole. All energies are relative to the 0-0 origin. Frequencies are in cm^{-1}. Experimental frequencies that have been extracted from the plots of REMPI spectra reported in Ref. [160] are marked by asterisk.

Based on such good accuracy of the simulated spectra, we can therefore attempt to suggest

[j] Experimental data from reference [160].
[k] In parenthesis the fundamental modes of S_1 resulting from the Duschinsky rotation between S_0 vibrations.
[m] Bands considered for the estimation of the RMS deviation.
[o] Average of four bands considered for the estimation of the RMS deviation.
[p] Alternative assignment: $\{0.67\ v'_{9a}\}$
[q] Alternative assignment: $\{\text{-}0.61\ v'_{6b}$ - $0.39\ v'_{6a}\}+\{\text{-}0.70\ v'_{12}\}$ combination.
[r] Average of 1696 cm^{-1} and 1713 cm^{-1} considered for the estimation of the RMS deviation.

assignments for the reported experimental vibronic transitions [159, 160]. Expansions of the most relevant spectral regions are shown in Figures 4.27 to 4.31, while the observed vibronic transitions are compared to the computed ones in Table 4.13 [with either the calculated / theoretical (TA), or the empirical / experimental (EA) anharmonic correction], which also lists the proposed assignment. The RMS deviation computed on all bands that have been assigned in this work amounts to 18 cm^{-1} and 15 cm^{-1}, for TA and EA anharmonic corrections, respectively.

The band at 234 cm^{-1}, assigned to a $\{v'_{\text{COC torsion}}\}$ + $\{v'_{10b}\}$ combination by Matsumoto et al. [159], is indeed displayed at 204 cm^{-1}. The lower intensity of the calculated band is in keeping with the proposed involvement of a Fermi resonance with $v'_{\text{COC bending}}$, an effect which is not accounted for at the present computational level. In correspondence with the band at 259 cm^{-1}, the simulated spectrum displays a peak of similar intensity (249 cm^{-1} for TA, 256 cm^{-1} for EA), which is connected to mode 5 of S_1 ($v'_{\text{COC bending}}$): this is in line with the assignment of Matsumoto et al. [159], but at variance with the proposal (v'_{18b}) by Hoffman et al. [160]. The band at 367 cm^{-1} is assigned in the literature to v'_{16b}; while this agrees with the computed frequency of this mode (376 cm^{-1} TA, 372 cm^{-1} EA), the simulated spectrum does not display any transition related to this fundamental. A seemingly plausible match at 340 cm^{-1} corresponds in fact to third overtone of v'_{16a}, which, like the first overtone (see section 4.8.3), needs to be considered as an artifact. This leaves as the best candidate the weaker transition at 384 cm^{-1}, that corresponds to the first overtone of mode 4 of S_1, resulting from Duschinsky rotation between $v'_{\text{O-CH}_3\text{ torsion}}$ and v'_{10b}; a similar interpretation has been proposed by Matsumoto based on the analysis of dispersed fluorescence (DF) spectra. The band at 427 cm^{-1}, assigned to v'_{15} by Matsumoto et al. [159] and to the *COC bending* by Hoffman et al. [160], can instead be correlated with the 7th (428 cm^{-1} TA, 421 cm^{-1} EA) or 8th (433 cm^{-1} TA, 436 cm^{-1} EA) mode of S_1. As the former is more intense, assignment to v'_{9b} is proposed.

In the region between 500 and 530 cm^{-1}, four bands have been observed and assigned to modes v'_{6a}, v'_{10a}, v'_4 and v'_{6b}. The simulated spectrum does show four transitions in this range: of these, the two most intense bands, mode 9 of S_1 (512 cm^{-1} TA, 509 cm^{-1} EA) and 11 of S_1 (540 cm^{-1} TA, 538 cm^{-1} EA), can be correlated to mixing of v'_{6b} and v'_{6a}, while the two weaker ones appear to be combinations of S_1 modes. A very weak transition at 621 cm^{-1} has been assigned to the v'_{11} fundamental, in line with the calculated frequency for this mode (623 cm^{-1} TA, 620 cm^{-1} EA). However, all the transitions in the 550-730 cm^{-1} range of the simulated spectrum are related to combinations of S_1 vibrations. Therefore, also the very weak band at 667 cm^{-1}, which has been assigned to mode v'_{17a}, and would match well the computed frequency of mode v'_{17b} (673 cm^{-1} TA, 670 cm^{-1} EA), should instead be regarded as a combination. An analogous situation concerns the band at 704 cm^{-1}. The intense peak at 759 cm^{-1}, for which an assignment to v'_1 is proposed in the literature, shows two additional weak features on its low-frequency shoulder (see panel A of Figure 4.27). Only the band at 750 cm^{-1} has been considered in earlier analyses, and assigned to v'_5; however, inspection of the simulated spectrum would rather suggest assignment to the first overtone of the 6th mode of S_1 ($\{$-0.49 v'_{16b} + 0.46 $v'_4\}$, 752 cm^{-1} TA, 744 cm^{-1} EA) or to the $\{v'_{\text{COC bending}}\}$ + $\{$-0.61 v'_{6b} - 0.39 $v'_{6a}\}$ combination (761 cm^{-1} TA, 765 cm^{-1} EA).

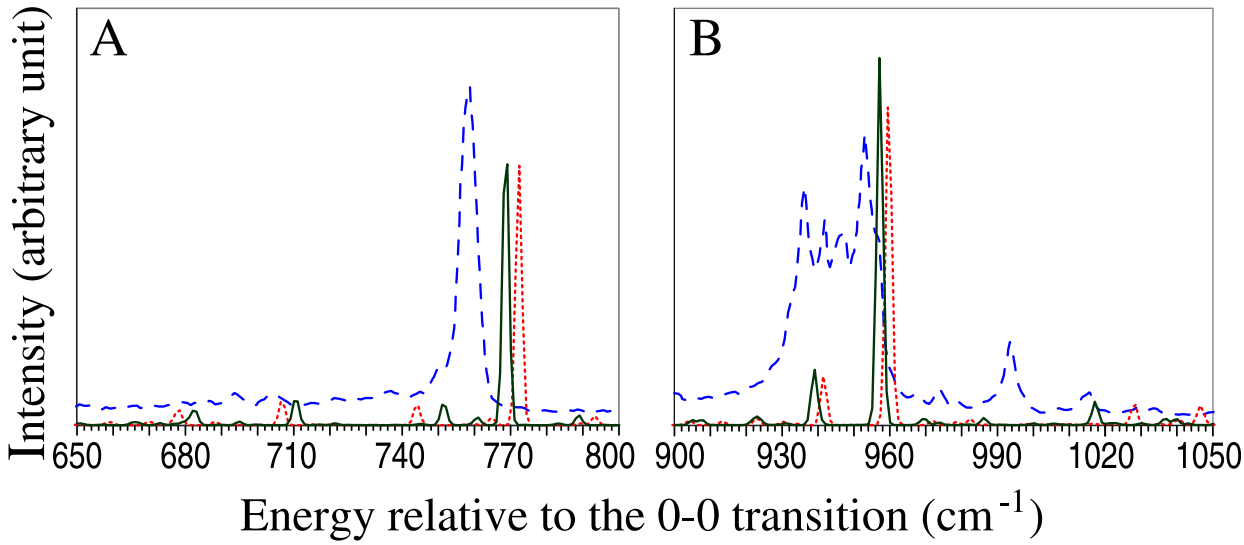


Figure 4.27: Theoretical and experimental [160] spectra of the $S_1 \leftarrow S_0$ transition of anisole from 590 cm^{-1} to 800 cm^{-1} and from 900 cm^{-1} to 1050 cm^{-1}. Green solid line, spectrum computed with inclusion of the "TA" anharmonic corrections; red dashed line, spectrum computed with the "EA" anharmonic correction; blue long-dashed line, experimental REMPI spectrum.

REMPI spectra display a group of four bands (see panel B of Figure 4.27) around 940 cm^{-1}, which in the literaure have been assigned to modes v'_{12}, v'_{15}, v'_{9a} and v'_{18a}. Only two intense peaks appear in this region of the simulated spectrum: the band at 957 cm^{-1} corresponds to the v'_{12} fundamental transition (mode 17 of S_1), while all others are combination bands (e.g. involving the 15th mode of S_1 (v'_{17a}). Moreover, in some cases the experimental intensity could reflect Fermi resonance with the v'_{12} band, a phenomenon which cannot be reproduced at the current level of approximation. Other ambiguities concern the band at 994 cm^{-1}, for which assignments as v'_{17a}, v'_{9b} and v'_{18b} have been proposed by Balfour [166], Hoffman et al. [160] and Matsumoto et al. [159], respectively. Of the three possibilities, analysis of the simulated spectra supports the latter one: the transition corresponding to the 19th mode of S_1 is much weaker, and can be described as an almost half and half mixing of v'_{18b} and $v'_{\text{O-CH}_3\text{ stretch.}}$. Concerning the band at 1016 cm^{-1}, assigned to *O-CH_3 stretching*, the computed frequency of mode 20 $\{0.52\ v'_{18b} + 0.44\ v'_{\text{O-CH}_3\text{ stretch.}}\}$ corresponds closely to the band position. However, if transition intensities are taken into account, assignment to a combination involving the v'_{17a} mode becomes preferable.

Both experimental and simulated spectra show many weak features between 1100 and 1350 cm^{-1} (see Figure 4.28). The bands at 1126, 1152 and 1179 cm^{-1} have been assigned to modes v'_3 and $v_{\text{CH}_3\text{ rock.}}$, respectively. Inspection of the simulated spectrum allows to correlate the strongest band (1127 cm^{-1} TA, 1128 cm^{-1} EA) to the combination involving v'_{12}.

The very weak, broad band at $\approx$ 1140 cm^{-1} (not included in the analysis by Hoffman et al.) can be correlated either with the weak transitions at 1136 cm^{-1} (1127 cm^{-1} EA; $\{0.55\ v'_{15} + 0.30\ v'_{9a}\}$), or with the 1143 cm^{-1} band (1136 cm^{-1} EA; $\{0.67\ v'_{9a}\}$). Assignment of the very weak band at 1152 cm^{-1} as a *CH_3 rock.* is confirmed by the simulation, while the experimental band at 1179 cm^{-1} can be associated in the computed spectrum to a peak at 1191 cm^{-1} (1189 cm^{-1} EA; first overtone of v'_{17a}), or else to the combination band at the 1205 cm^{-1}. An assignment to

mode v'_{8a} has been proposed for the signal at 1271 cm^{-1}; however, contrary to CIS calculations, our results do not support such a strong red-shift of this mode. We propose instead assignment to the v'_{7b} fundamental which appears in simulated spectrum at 1247 cm^{-1} (1258 cm^{-1} EA). The band at 1288 cm^{-1} has been assigned to v'_{13} [159] or to v'_{7a} [160]; by contrast, the simulated spectrum features a transition at 1284 cm^{-1} (1277 cm^{-1} EA), which is correlated to the $\{0.56\ v'_3 + 0.30\ v'_{14}\}$ fundamental, and gains intensity from overlap with a strong combination band.

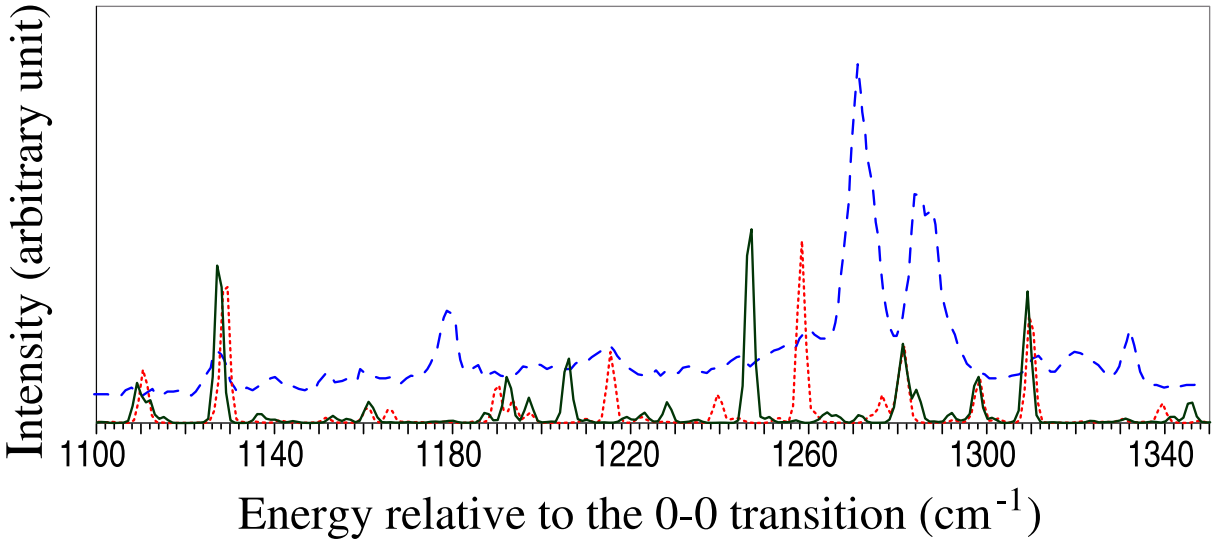


Figure 4.28: (Theoretical and experimental [160] spectra of the $S_1 \leftarrow S_0$ transition of anisole from 1100 cm^{-1} to 1350 cm^{-1}. Refer to legend of Figure 4.27 for details.

The difficulty of performing assignments in the spectral range between 1400 and 1500 cm^{-1} (see Figure 4.29) has already been pointed out by Hoffman et al. [160].

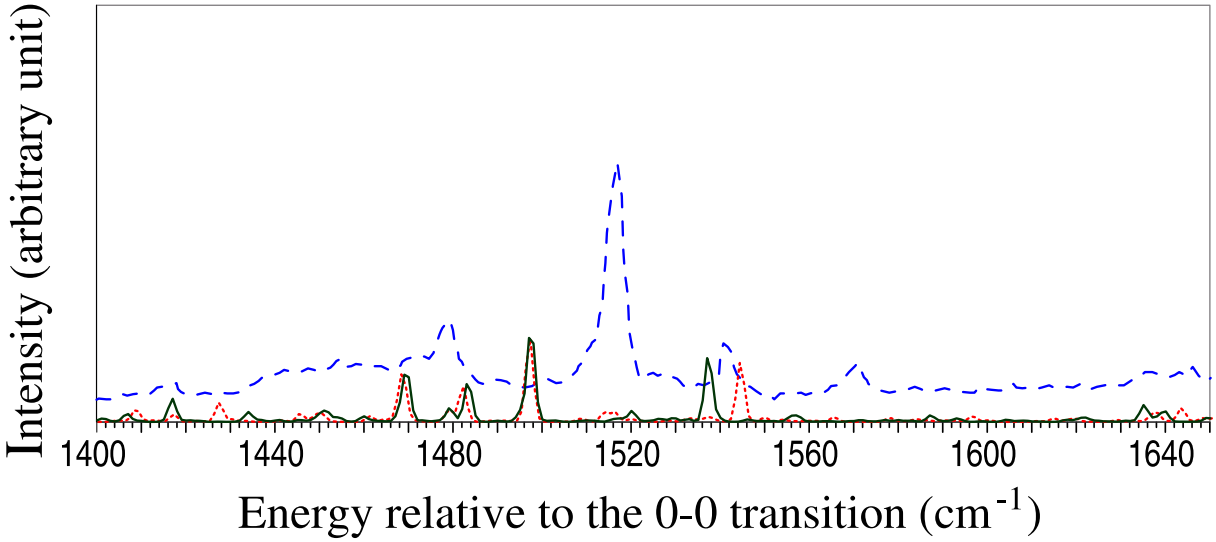


Figure 4.29: Theoretical and experimental [160] spectra of the $S_1 \leftarrow S_0$ transition of anisole 1400 cm^{-1} to 1650 cm^{-1}. Refer to legend of Figure 4.27 for details.

Based on the relative intensities in the simulated spectrum, we suggest that most of the bands

in the 1400-1460 region that were reported [160] as fundamental transitions, should rather be interpreted as combinations. Only the one at 1443 cm^{-1} can be related to a fundamental transition, namely mode 29 of S_1 $\{0.66\ v'_{CH_3\ sym.\ def.} + 0.21\ v'_{19a}\}$ (1433 cm^{-1} TA, 1417 cm^{-1} EA). Conversely, the simulated spectrum correlates the S_1 fundamentals v'_{19b} (1390 cm^{-1} TA, 1389 cm^{-1} EA), v'_{8a} (1423 cm^{-1} TA, 1424 cm^{-1} EA), v'_{19a} (1450 cm^{-1} TA, 1446 cm^{-1} EA), and $v'_{CH_3\ antisym.\ def.}$ (1458 cm^{-1} TA, 1454 cm^{-1} EA) to transitions of very low intensity. For the band at 1479 cm^{-1}, reassignment from v'_{19b} to $\{-0.81\ v'_{8b}\}$ or $\{-0.61\ v'_{6b} - 0.39\ v'_{6a}\} + \{-0.77\ v'_{11}\}$ is proposed, in consideration of the presence of peaks of compatible intensity at 1483 cm^{-1} (1482 cm^{-1} EA) and 1469 cm^{-1} (both TA and EA), respectively. The strong transition at 1517 cm^{-1} has been recently assigned by Hoffman et al. as mode v'_{19a} involved in Fermi interaction with the first overtone of v'_1; however, in the simulated spectrum the first overtone of v'_1 appears at 1536 cm^{-1} (1544 cm^{-1} EA), and matches the experimental band at about 1540 cm^{-1}, not analysed in [160]. In turn the band at 1517 cm^{-1}, should be instead assigned as the $\{v'_{12}\} + \{-0.60\ v'_{6a} + 0.38\ v'_{6b}\}$ combination, related to an intense transition at 1497 cm^{-1}. Weak bands at 1571 and 1636 cm^{-1} have been assigned as v'_{8a} and v'_{14} fundamentals, an assignment that is not supported by the present results, since the simulated spectrum associates to those bands only weak combination transitions.

In a range between 1650 and 1750 cm^{-1} (see Figure 4.30), the REMPI spectrum shows a group of intense bands, which have been assigned as combinations involving the v'_1 mode [160].

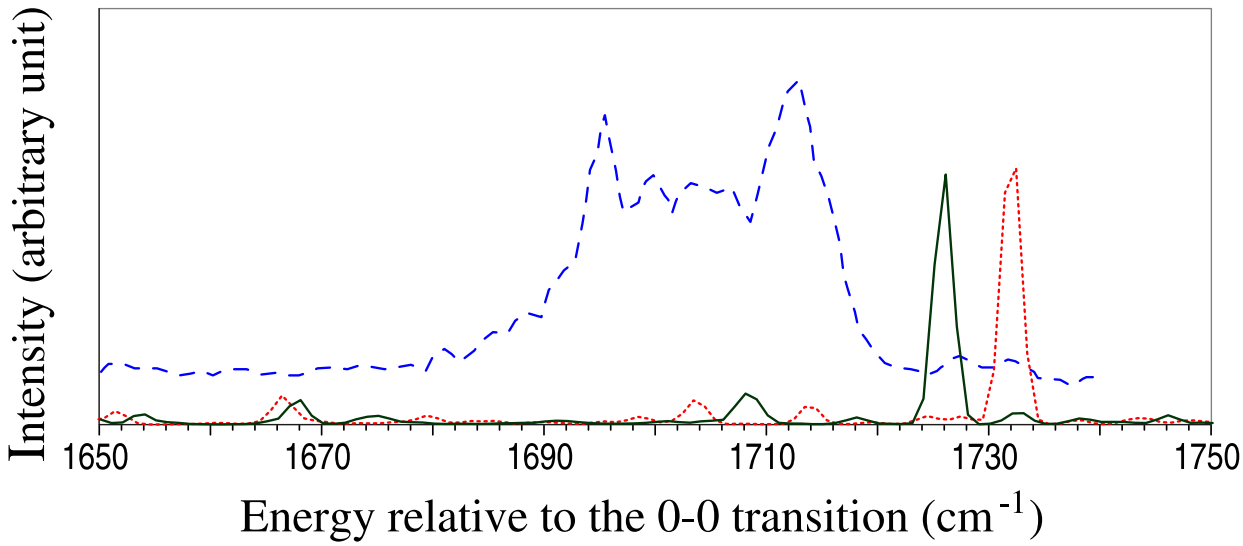


Figure 4.30: Theoretical and experimental [160] spectra of the $S_1 \leftarrow S_0$ transition of anisole from from 1650 cm^{-1} to 1750 cm^{-1}. Refer to legend of Figure 4.27 for details.

These findings are in line with our results, although most computed transitions appear less intense than their experimental counterparts. The part of the spectrum above 1800 cm^{-1} displays a rich structure of low intensity bands, most of which are actually related to overtones and combinations. Thus, any attempt to assign transitions to the high frequency fundamentals of S_1 should be considered as tentative. Bands at 2967, 2979, and 2990 cm^{-1} have been assigned to the symmetric and antisymmetric *C-H stretching* vibrations of CH_3. The empirical anharmonic frequencies (2928, 2980, and 3025 cm^{-1}, respectively), agree well with the above mentioned experimental data, but a specific transition can only be associated to the $v'_{CH_3\ sym.\ stretch.}$ (as a matter of fact, a very weak

one). The spectral region related to the C-X frequencies of the aromatic ring is shown on Figure 4.31.

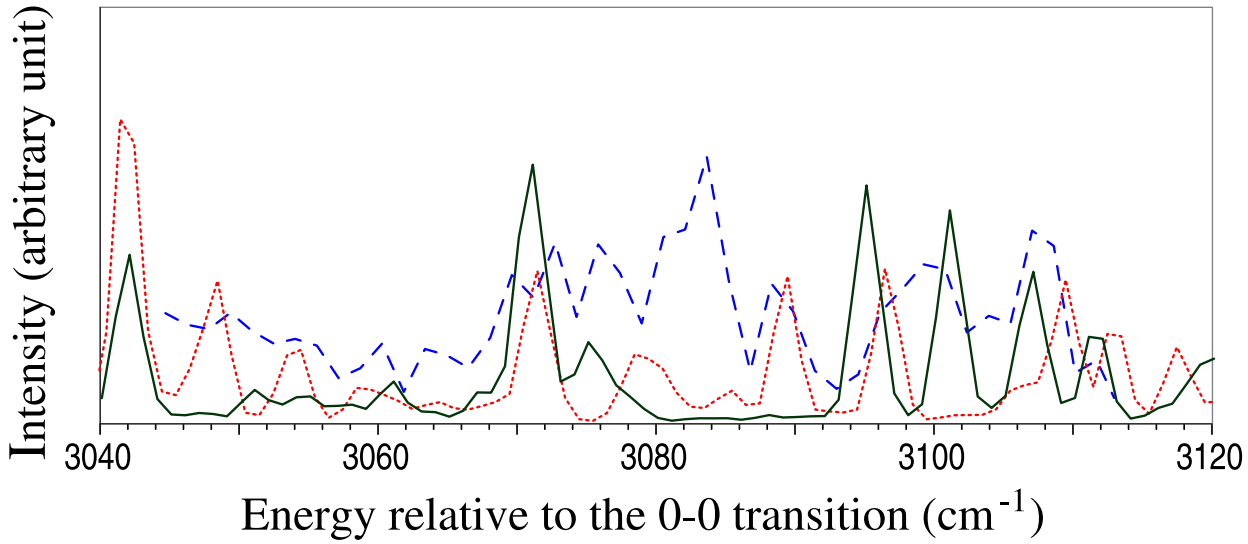


Figure 4.31: Theoretical and experimental [160] spectra of the $S_1 \leftarrow S_0$ transition of anisole from 3040 cm^{-1} to 3120 cm^{-1}. Refer to legend of Figure 4.27 for details.

It is apparent that there are many possible bands which could be assigned to the S_1 fundamental transitions. The simulated spectrum displays very weak transitions related to v'_{13}, v'_{20a}, v'_{20b}, and v'_2, at 3024, 3061, 3095, and 3101 cm^{-1}, respectively (TA). For v'_{13} and v'_{20a}, band intensity is partially gained from the interaction with overtones or combinatios involving the v'_1 and v'_{12} modes. It should be noted that no transition can be associated to the v'_{7a} fundamental, at variance with the assignment of the band at 3084 cm^{-1}.

On the whole, comparison of the experimental REMPI spectrum with its computed counterpart, which we have explored in the present work, has provided several new insights on the assignment of the vibronic bands of the $S_1 \leftarrow S_0$ electronic transition of anisole. In particular, for many bands that had been assigned to S_1 fundamentals, consideration of the relative intensities has suggested instead a different interpretation, as combinations or overtones.

Conclusions

The procedure presented in this thesis has been used here for an extensive study of a practical case, anisole. On the whole, a very good agreement between computed and experimental vibrationally resolved REMPI spectrum has been achieved, which confirms the good accuracy of the DFT/TD-DFT geometries and force fields computed for the ground and excited electronic state, respectively. In order to reproduce correctly the band intensities and the rich vibrational structure of the REMPI spectra, it has been necessary to account for changes in structure, vibrational frequencies and normal modes between the involved electronic states. It is worthwhile underlining that the remarkable overall agreement, also as far as band positions are concerned, has only been possible when the frequencies have been corrected for anharmonicity. This implies that the simple scheme proposed in the present work can be effectively applied to derive mode specific scaling factors for molecular systems in excited electronic states. In particular, the empirical anharmonic

correction (EA) benefits from the availability of accurate and well-defined ground state data, and thus can be effectively exploited even for large systems. The discrepancy between the absolute position of experimental and simulated spectra needs to be mentioned as the main shortcoming of the purely theoretical approach. To achieve an appropriate fit between spectra, the energy of the electronic transition would need to be computed with the accuracy of ≈ 10 cm^{-1}. Thus, even if DFT/TD-DFT computations are able to provide quite reasonable estimates of the relative energetics of the electronic states, and despite the adoption of a refinement based on Coupled Clusters calculations, it was still necessary to compare spectra shifted to the 0-0 origin.

With this caveat, the computational approach proposed in the present work leads to a remarkable agreement with high-accuracy spectroscopic data. In conclusion, it seems evident that the approach we adopted here is more reliable that a comparison based purely on computed frequencies, and represents a valuable tool for the interpretation of experimental results.

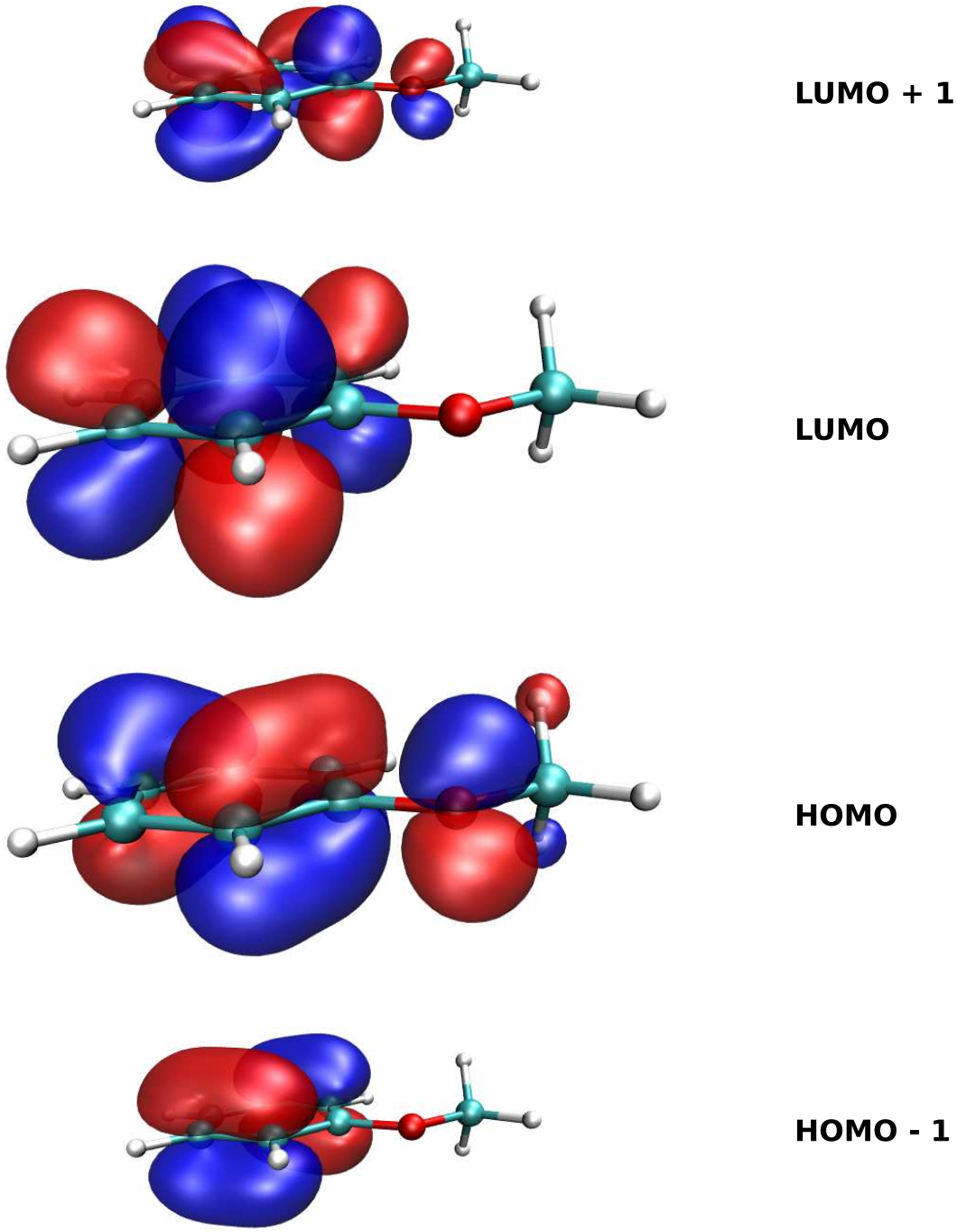


Figure 4.32: Frontier orbitals involved in the $S_1 \leftarrow S_0$ transition of anisole, calculated at the TD-B3LYP/6-311+G(d,p) level for the geometry of the final state.

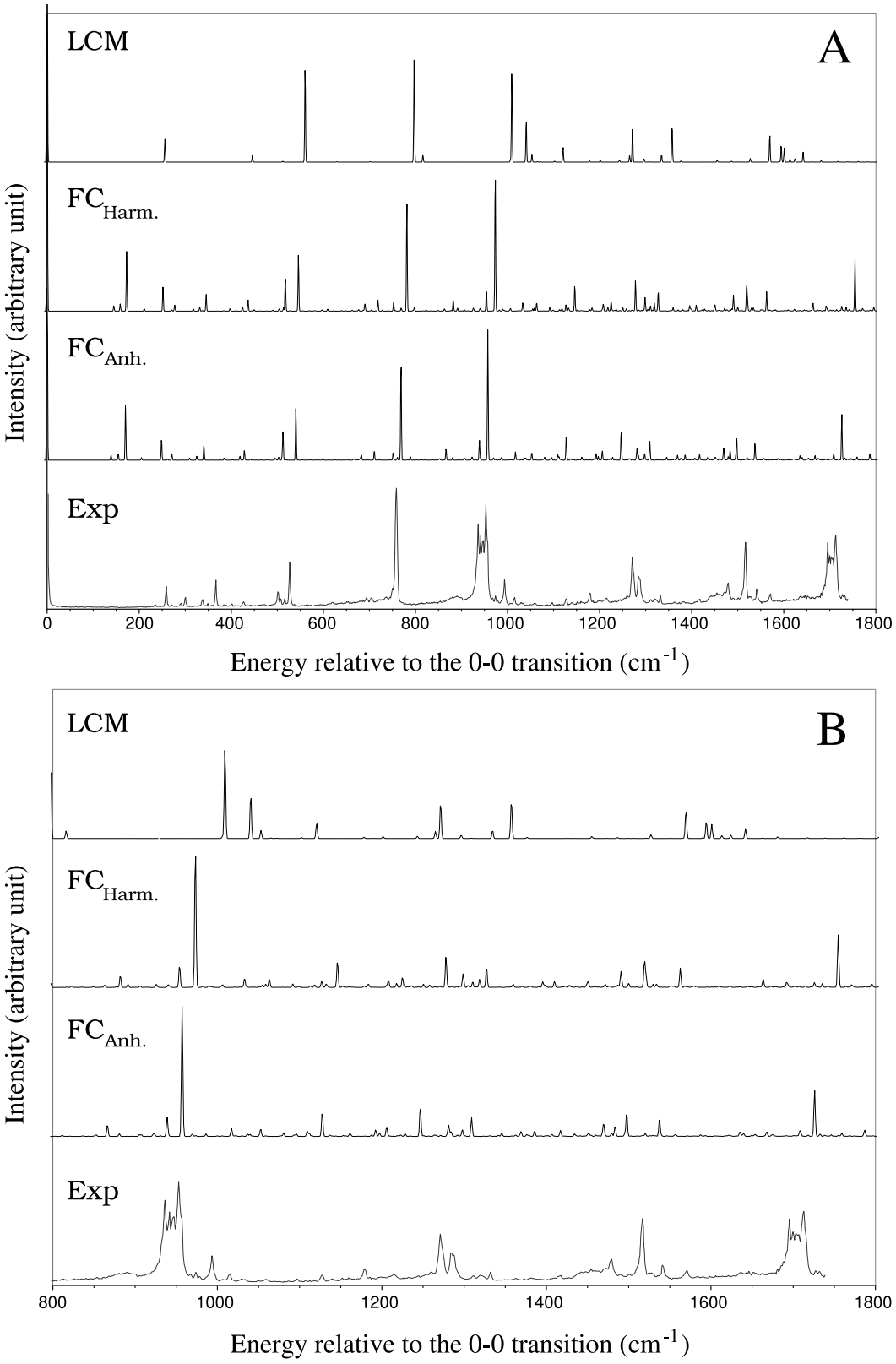


Figure 4.33: Theoretical and experimental [160] spectra of the $S_1 \leftarrow S_0$ transition of anisole. Spectra computed within the LCM [LCM], Harmonic Franck-Condon (including Duschinsky rotation) [$FC_{Harm.}$], and Anharmonic Franck-Condon [$FC_{Anh.}$] approximations are shown. Panel A, 0-1800 cm^{-1} range; panel B, expansion of the 800-1800 cm^{-1} region.

	S_0				S_1	
	Ref. [158][s)]	Ref. [159][t)]	Ref. [167][u)]	This work[v)]	Ref. [159][w)]	This work[x)]
C1C2	1.400	1.401	1.404	1.401	1.422	1.424
C2C3	1.388	1.387	1.393	1.388	1.410	1.425
C3C4	1.398	1.397	1.402	1.398	1.405	1.412
C4C5	1.390	1.390	1.395	1.390	1.415	1.418
C5C6	1.398	1.397	1.402	1.398	1.410	1.426
C6C1	1.398	1.397	1.402	1.397	1.417	1.421
C1O7	1.366	1.365	1.364	1.366	1.321	1.349
O7C8	1.421	1.420	1.418	1.420	1.407	1.429
C2H9	1.083	1.083		1.083	1.072	1.081
C3H10	1.084	1.084		1.084	1.073	1.081
C4H11	1.084	1.084		1.083	1.075	1.085
C5H12	1.085	1.085		1.084	1.073	1.081
C6H13	1.082	1.082		1.082	1.070	1.080
C8H14(16)	1.096	1.096		1.096	1.084	1.093
C8H15	1.089	1.089		1.089	1.079	1.088
C1O7C8	118.6	118.5	116.4	118.6	121.4	121.0
C2C1O7	115.7	115.7		115.7	114.6	113.9
C6C1O7	124.5	124.6	124.9	124.5	123.5	122.9
C1C2C3	120.4	120.1	121.0	120.0	119.1	118.1
C2C3C4	120.6	120.5	119.4	120.6	119.2	119.1
C3C4C5	119.2	119.2	121.0	119.2	121.5	122.4
C4C5C6	120.9	121.0	119.2	120.9	120.0	119.4
C5C6C1	119.5	119.5		119.5	118.2	117.7
C6C1C2	119.8	119.7		119.8	121.9	123.2
H15C8O7				105.9		105.5
H14C8O7				111.4		110.6
H15C8O7C1				180.0		180.0
H14C8O7C2				61.3		61.0
H16C8O7C2				-61.3		-61.0

Table 4.14: Calculated geometrical parameters (bond lengths in Å and angles in degrees) of anisole in its ground and excited electronic state.

[s]B3LYP / 6-311++G(d,p)
[t]B3LYP / 6-311G(d,p)
[u]MP2 / 6-311G(d,p)
[v]B3LYP / 6-311+G(d,p)
[w]CIS / 6-311G(d,p)
[x]TD-B3LYP / 6-311+G(d,p)

mode	CIS/6-31G(d,p)[y)]	CIS/6-311G(d,p)[z)]	TD/6-311+G(d,p)		
			ω''	$\vartheta''_{\mathrm{TA}}$	$\vartheta''_{\mathrm{EA}}$
1	88	83	72	69	66
2	162	164	86	85	85
3	172	182	139	136	140
4	221	225	199	192	197
5	247	247	252	249	256
6	362	360	376	376	372
7	418	416	437	428	421
8	499	498	441	433	436
9	523	512	517	512	509
10	526	520	530	529	524
11	529	524	546	540	538
12	634	625	604	596	595
13	671	655	629	623	620
14	713	711	680	673	670
15	757	757	781	769	772
16	841	838	858	850	845
17	944	947	973	957	960
18	961	964	990	970	972
19	998	1004	1006	986	982
20	1036	1044	1035	1014	1008
21	1076	1092	1143	1121	1118
22	1133	1135	1152	1136	1127
23	1154	1149	1161	1144	1135
24	1179	1179	1183	1161	1161
25	1234	1247	1278	1247	1258
26	1288	1290	1310	1284	1277
27	1333	1343	1395	1369	1364
28	1415	1419	1420	1390	1389
29	1424	1428	1449	1434	1417
30	1463	1464	1457	1423	1424
31	1465	1466	1476	1450	1446
32	1474	1476	1483	1470	1443
33	1561	1571	1494	1458	1454
34	1629	1642	1521	1483	1482
35	2880	2878	3032	2897	2928

Table 4.15: Calculated frequencies (in cm^{-1}) of anisole in its first excited electronic state.

[y]Ref. [160] CIS/6-31G(d,p) harmonic frequencies scaled by a factor of 0.9
[z]Ref. [159] CIS/6-311G(d,p) harmonic frequencies scaled by a factor of 0.905

Table 4.15: Continuation

mode	CIS/6-31G(d,p)[y)]	CIS/6-311G(d,p)[z)]	TD/6-311+G(d,p)		
			ω''	$\vartheta''_{\mathrm{TA}}$	$\vartheta''_{\mathrm{EA}}$
36	2944	2945	3101	2940	2980
37	2986	2984	3153	3014	3025
38	3020	3021	3171	3024	3042
39	3036	3038	3201	3072	3066
40	3046	3048	3208	3061	3078
41	3060	3064	3223	3095	3117
42	3071	3074	3229	3101	3128

Table 4.15: Calculated frequencies (in cm^{-1}) of anisole in its first excited electronic state.

[y]Ref. [160] CIS/6-31G(d,p) harmonic frequencies scaled by a factor of 0.9
[z]Ref. [159] CIS/6-311G(d,p) harmonic frequencies scaled by a factor of 0.905

Conclusion

In this document, we presented a complete approach to generate vibrationally resolved electronic spectra. The implementation within a quantum chemical calculation package such as GAUSSIAN [18] raises some issues but offers the possibility of a versatile and fully integrated approach. Thanks to this, it is possible to carry out a complete calculation starting from the geometry optimization and up to the simulation of the spectrum as requested by the user. Since we intended our procedure to be particularly efficient for medium-to-large systems such as molecules of biological interest, a particular care has been devoted to the theoretical method underlying the calculation of the spectrum and its encoding.

Finally, we presented several applications of our method ranging from illustrations of the possibilities it offers to extended examples as a support to analyze experimental data. While our procedure already gives accurate results, ways of improvement still remain. We already mentioned some further developments which are currently under way such as the inclusion of the temperature and the possibilities to refer to the initial state for the Taylor expansion of the electronic transition dipole moment. It should also be noted that in some cases, the prescreening method is not able to attain a fully converged spectrum with respect to the analytic sum rules. We saw previously that this could be caused by setting $N_I^{\max}$ too low with respect to the number of normal modes of the system. However, it can also happen that the chosen vector $\boldsymbol{v}''_{\max}$ is not well suited, leading in this case to an over-estimation of the limit of the number of transitions to compute in each class [168]. Improvements in the evaluation of the transitions to handle would help to reduce the number of transitions to compute and so to speed up the calculations when dealing with large systems. Nevertheless, as noted in the case of chlorophyll *c2*, the shape convergence of the spectrum is faster than the sum of the squared transition dipole moment integrals. Consequently, the procedure can be effectively used to analyze electronic spectra even when the convergence with respect to the analytic sum rule is significantly lower than 100%.

Therefore, the current procedure already provides a fully functional framework to simulate the electronic spectra of a large range of systems. Moreover, while designed to be easily accessible to non-specialists, it offers several parameters to control the calculations. As a result, it can be adapted to very specific needs or to match diverse experimental conditions.

Appendix A

General formula of the transition dipole moment integral

Using the boson creation and annihilation operators, the transition dipole moment integral can be written

$$
\begin{aligned}
\langle\, \Psi' \mid \boldsymbol{\mu} \mid \Psi'' \,\rangle =& \boldsymbol{\mu}_{if}(\mathbf{Q}_0'') \,\langle\, \boldsymbol{v}' \mid \boldsymbol{v}'' \,\rangle \\
&+ \sum_{k=1}^{N} \left(\frac{\partial \boldsymbol{\mu}_{if}}{\partial Q_k''} \right)_0 \sqrt{\frac{\hbar}{2\omega_i''}} \left[\sqrt{v_k''} \,\langle\, \boldsymbol{v}' \mid \boldsymbol{v}'' - 1_k'' \,\rangle \right. \\
&\quad \left. + \sqrt{v_k'' + 1} \,\langle\, \boldsymbol{v}' \mid \boldsymbol{v}'' + 1_k'' \,\rangle \right] \\
&+ \sum_{k=1}^{N} \left(\frac{\partial^2 \boldsymbol{\mu}_{if}}{\partial Q_k''^2} \right)_0 \frac{\hbar}{4\omega_k''} \left[\sqrt{v_k''(v_k'' - 1)} \,\langle\, \boldsymbol{v}' \mid \boldsymbol{v}'' - 2_k'' \,\rangle + (2v_k'' + 1) \,\langle\, \boldsymbol{v}' \mid \boldsymbol{v}'' \,\rangle \right. \\
&\quad \left. + \sqrt{(v_k'' + 1)(v_k'' + 2)} \,\langle\, \boldsymbol{v}' \mid \boldsymbol{v}'' + 2_k'' \,\rangle \right] \\
&+ \sum_{k=1}^{N} \sum_{\substack{l=1 \\ l \neq k}}^{N} \left(\frac{\partial^2 \boldsymbol{\mu}_{if}}{\partial Q_k'' \partial Q_l''} \right)_0 \frac{\hbar}{4\left(\sqrt{\omega}_k'' \sqrt{\omega}_l''\right)} \left[\sqrt{v_k'' v_l''} \,\langle\, \boldsymbol{v}' \mid \boldsymbol{v}'' - 1_k'' - 1_l'' \,\rangle \right. \\
&+ \sqrt{v_k''(v_l'' + 1)} \,\langle\, \boldsymbol{v}' \mid \boldsymbol{v}'' - 1_k'' + 1_l'' \,\rangle + \sqrt{(v_k'' + 1) v_l''} \,\langle\, \boldsymbol{v}' \mid \boldsymbol{v}'' + 1_k'' - 1_l'' \,\rangle \\
&\left. + \sqrt{(v_k'' + 1)(v_l'' + 1)} \,\langle\, \boldsymbol{v}' \mid \boldsymbol{v}' + 1_k'' + 1_l'' \,\rangle \right]
\end{aligned}
\tag{A.1}
$$

By reordering the terms according to the overlap integrals, it is possible to rewrite the previous equation in the following, more convenient form:

$$
\begin{aligned}
\langle\, \Psi' \mid \boldsymbol{\mu} \mid \Psi'' \,\rangle = & \left[\boldsymbol{\mu}_{if}(\mathbf{Q}_0'') + \sum_{k=1}^{N} \frac{\hbar}{4\omega_k''} \left(\frac{\partial^2 \boldsymbol{\mu}_{if}}{\partial Q_k''^2} \right)_0 (2v_k'' + 1) \right] \langle\, \boldsymbol{v}' \mid \boldsymbol{v}'' \,\rangle \\
& + \sum_{k=1}^{N} \frac{\hbar}{4\omega_k''} \left(\frac{\partial^2 \boldsymbol{\mu}_{if}}{\partial Q_k''^2} \right)_0 \sqrt{v_k''(v_k'' - 1)}\, \langle\, \boldsymbol{v}' \mid \boldsymbol{v}'' - 2_k \,\rangle \\
& + \sum_{k=1}^{N} \left\{ \sqrt{\frac{\hbar}{2\omega_k''}} \left(\frac{\partial \boldsymbol{\mu}_{if}}{\partial Q_k''} \right)_0 \sqrt{v_k''}\, \langle\, \boldsymbol{v}' \mid \boldsymbol{v}'' - 1_k'' \,\rangle \right. \\
& \quad + \sum_{\substack{l=1 \\ l \neq k}}^{N} \frac{\hbar}{4\sqrt{\omega_k''\omega_l''}} \left(\frac{\partial^2 \boldsymbol{\mu}_{if}}{\partial Q_k'' \partial Q_l''} \right)_0 \left[\sqrt{v_k'' v_l''}\, \langle\, \boldsymbol{v}' \mid \boldsymbol{v}'' - 1_k'' - 1_l'' \,\rangle \right. \\
& \quad \left. \left. + \sqrt{v_k''(v_l'' + 1)}\, \langle\, \boldsymbol{v}' \mid \boldsymbol{v}'' - 1_k'' + 1_l'' \,\rangle \right] \right. \\
& \quad + \sqrt{\frac{\hbar}{2\omega_k''}} \left(\frac{\partial \boldsymbol{\mu}_{if}}{\partial Q_k''} \right)_0 \sqrt{v_k'' + 1}\, \langle\, \boldsymbol{v}' \mid \boldsymbol{v}'' + 1_k'' \,\rangle \\
& \quad + \sum_{\substack{l=1 \\ l \neq k}}^{N} \frac{\hbar}{4\sqrt{\omega_k''\omega_l''}} \left(\frac{\partial^2 \boldsymbol{\mu}_{if}}{\partial Q_k'' \partial Q_l''} \right)_0 \left[\sqrt{(v_k'' + 1) v_l''}\, \langle\, \boldsymbol{v}' \mid \boldsymbol{v}'' + 1_k'' - 1_l'' \,\rangle \right. \\
& \quad \left. \left. + \sqrt{(v_k'' + 1)(v_l'' + 1)}\, \langle\, \boldsymbol{v}' \mid \boldsymbol{v}'' + 1_k'' + 1_l'' \,\rangle \right] \right\} \\
& + \sum_{k=1}^{N} \frac{\hbar}{4\omega_k''} \left(\frac{\partial^2 \boldsymbol{\mu}_{if}}{\partial Q_k''^2} \right)_0 \sqrt{(v_k'' + 1)(v_k'' + 2)}\, \langle\, \boldsymbol{v}' \mid \boldsymbol{v}'' + 2_k \,\rangle
\end{aligned}
\tag{A.2}
$$

Appendix B

Details for the calculation of the overlap integral $\langle\, 0' \mid 0'' \,\rangle$

Let us consider the exponent of e in equation 1.71, $-\frac{1}{2}(\mathbf{Q'}^{\mathrm{T}}\mathbf{\Gamma'}\mathbf{Q'}+\mathbf{Q''}^{\mathrm{T}}\mathbf{\Gamma''}\mathbf{Q''})$. Using the Duschinsky transformation given in equation 1.64, this term can be expanded into the following equation:

$$\begin{aligned}\mathbf{Q'}^{\mathrm{T}}\mathbf{\Gamma'}\mathbf{Q'} + \mathbf{Q''}^{\mathrm{T}}\mathbf{\Gamma''}\mathbf{Q''} &= \left(\mathbf{J}\mathbf{Q''} + \mathbf{K}\right)^{\mathrm{T}}\mathbf{\Gamma'}(\mathbf{J}\mathbf{Q''} + \mathbf{K}) + \mathbf{Q''}^{\mathrm{T}}\mathbf{\Gamma''}\mathbf{Q''} \\ &= \mathbf{Q''}^{\mathrm{T}}\mathbf{J}^{\mathrm{T}}\mathbf{\Gamma'}\mathbf{J}\mathbf{Q''} + \mathbf{Q''}^{\mathrm{T}}\mathbf{J}^{\mathrm{T}}\mathbf{\Gamma'}\mathbf{K} + \mathbf{K}^{\mathrm{T}}\mathbf{\Gamma'}\mathbf{J}\mathbf{Q''} + \mathbf{K}^{\mathrm{T}}\mathbf{\Gamma'}\mathbf{K} + \mathbf{Q''}^{\mathrm{T}}\mathbf{\Gamma''}\mathbf{Q''}\end{aligned} \tag{B.1}$$

Hence, it means integrating on the whole configuration space the exponential of a trinomial which is not straightforward. To simplify the problem, we will choose a linear transformation to the normal coordinates vector $\mathbf{Q''}$ so that the above polynomial can be reduced to a binomial of the form $x^2 + b$ where b is a constant. In our matrix notation, it will mean to find a vector $\mathbf{a}$[a] satisfying the relation:

$$\mathbf{Q'}^{\mathrm{T}}\mathbf{\Gamma'}\mathbf{Q'} + \mathbf{Q''}^{\mathrm{T}}\mathbf{\Gamma''}\mathbf{Q''} = \mathbf{a}^{\mathrm{T}}\mathbf{a} + \mathbf{d} \tag{B.2}$$

The vector $\mathbf{a}$ is related to the normal coordinates $\mathbf{Q''}$ by a linear transformation:

$$\mathbf{a} = \mathbf{b}\mathbf{Q''} + \mathbf{c} \tag{B.3}$$

Now, we will expand the right-hand side of equation B.2 to identify $\mathbf{b}$, $\mathbf{c}$ and $\mathbf{d}$ using the development of the exponent of e we wrote earlier.

$$\mathbf{a}^{\mathrm{T}}\mathbf{a} = \left(\mathbf{b}\mathbf{Q''} + \mathbf{c}\right)^{\mathrm{T}}(\mathbf{b}\mathbf{Q''} + \mathbf{c}) \tag{B.4}$$

$$= \mathbf{Q''}^{\mathrm{T}}\mathbf{b}^{\mathrm{T}}\mathbf{b}\mathbf{Q''} + \mathbf{Q''}^{\mathrm{T}}\mathbf{b}^{\mathrm{T}}\mathbf{c} + \mathbf{c}^{\mathrm{T}}\mathbf{b}\mathbf{Q''} + \mathbf{c}^{\mathrm{T}}\mathbf{c} \tag{B.5}$$

[a] To avoid misinterpretation with the other matrices used throughout this document, we will temporarily break our typographic convention and use lower case characters for the matrices of the transformation.

Now, equating the "coefficients" of $\mathbf{Q}''$, we can find the following relations:

$$\begin{aligned}
&\mathbf{b}^{\mathrm{T}}\mathbf{b} = \mathbf{J}^{\mathrm{T}}\boldsymbol{\Gamma}'\mathbf{J} + \boldsymbol{\Gamma}'' \\
&\quad \Rightarrow \mathbf{b} = (\mathbf{J}^{\mathrm{T}}\boldsymbol{\Gamma}'\mathbf{J} + \boldsymbol{\Gamma}'')^{1/2} && \text{(B.6)} \\
&\mathbf{c}^{\mathrm{T}}\mathbf{b} = \mathbf{K}^{\mathrm{T}}\boldsymbol{\Gamma}'\mathbf{J} \\
&\quad \Rightarrow \mathbf{c} = (\mathbf{J}^{\mathrm{T}}\boldsymbol{\Gamma}'\mathbf{J} + \boldsymbol{\Gamma}'')^{-1/2}\mathbf{J}^{\mathrm{T}}\boldsymbol{\Gamma}'\mathbf{K} && \text{(B.7)} \\
&\mathbf{c}^{\mathrm{T}}\mathbf{c} + \mathbf{d} = \mathbf{K}^{\mathrm{T}}\boldsymbol{\Gamma}'\mathbf{K} \\
&\quad \Rightarrow \mathbf{d} = \mathbf{K}^{\mathrm{T}}\boldsymbol{\Gamma}'\mathbf{K} - \mathbf{K}^{\mathrm{T}}\boldsymbol{\Gamma}'\mathbf{J}(\mathbf{J}^{\mathrm{T}}\boldsymbol{\Gamma}'\mathbf{J} + \boldsymbol{\Gamma}'')^{-1}\mathbf{J}^{\mathrm{T}}\boldsymbol{\Gamma}'\mathbf{K} && \text{(B.8)}
\end{aligned}$$

It is noteworthy that $\mathbf{b}$ is symmetric. This will prove to be useful when dealing with the formalism for polyatomic molecules.

To perform the integration, it is necessary to change the variables of integrations from $\{Q_i''\}$ to $\{\mathrm{a}_i\}$. To do so, we need to calculate the Jacobian determinant whose elements are given by the relation:

$$\frac{\partial Q_i''}{\partial \mathrm{a}_j} = \frac{\partial \sum_{k=1}^{N} (\mathrm{b}^{-1})_{ik}(\mathrm{a}_k - \mathrm{c}_k)}{\partial \mathrm{a}_j} = (\mathrm{b}^{-1})_{ik} \tag{B.9}$$

As a result, the Jacobian determinant is the determinant of the inverse matrix of $\mathbf{b}$.

Appendix C

Calculation of the power series $f(\mathbf{T}, \mathbf{U})$

Let us rewrite equation 2.6 so that only terms in the exponential depending on $\mathbf{Q}''$ remains in the integral.

$$\begin{aligned} f(\mathbf{T},\mathbf{U}) = \pi^{-N/2} \det\big[\mathbf{\Gamma}'\,\mathbf{\Gamma}''\big]^{1/4} \det(\mathbf{J})^{1/2} \exp\big[-\mathbf{U}^2 - \mathbf{T}^2 - \frac{1}{2}\mathbf{K}^{\mathrm{T}}\mathbf{\Gamma}'\mathbf{K} + 2\mathbf{T}^{\mathrm{T}}\mathbf{\Gamma}'^{1/2}\mathbf{K}\big] \\ \int d\mathbf{Q}'' \exp\Big[-\frac{1}{2}\big(\mathbf{Q}''^{\mathrm{T}}\mathbf{J}^{\mathrm{T}}\mathbf{\Gamma}'\mathbf{J}\mathbf{Q}'' + \mathbf{Q}''^{\mathrm{T}}\mathbf{J}^{\mathrm{T}}\mathbf{\Gamma}'\mathbf{K} + \mathbf{K}^{\mathrm{T}}\mathbf{\Gamma}'\mathbf{J}\mathbf{Q}'' \\ + \mathbf{Q}''^{\mathrm{T}}\mathbf{\Gamma}''\mathbf{Q}'' - 4\mathbf{U}^{\mathrm{T}}\mathbf{\Gamma}''^{1/2}\mathbf{Q}'' - 4\mathbf{T}^{\mathrm{T}}\mathbf{\Gamma}'^{1/2}\mathbf{J}\mathbf{Q}''\big)\Big] \end{aligned} \tag{C.1}$$

Let us apply the same change of variable as in equation B.3 so that the following relation is true:

$$\begin{aligned} \mathbf{a}^{\mathrm{T}}\mathbf{a} + \mathbf{d} = \mathbf{Q}''^{\mathrm{T}}\mathbf{J}^{\mathrm{T}}\mathbf{\Gamma}'\mathbf{J}\mathbf{Q}'' + \mathbf{Q}''^{\mathrm{T}}\mathbf{J}^{\mathrm{T}}\mathbf{\Gamma}'\mathbf{K} + \mathbf{K}^{\mathrm{T}}\mathbf{\Gamma}'\mathbf{J}\mathbf{Q}'' + \\ \mathbf{Q}''^{\mathrm{T}}\mathbf{\Gamma}''\mathbf{Q}'' - 4\mathbf{U}^{\mathrm{T}}\mathbf{\Gamma}''^{1/2}\mathbf{Q}'' - 4\mathbf{T}^{\mathrm{T}}\mathbf{\Gamma}'^{1/2}\mathbf{J}\mathbf{Q}'' \end{aligned} \tag{C.2}$$

The development on the right side of equation is given in equation B.5. Equating the coefficients of $\mathbf{Q}''$, we find that $\mathbf{b}$ has the same value as in equation B.6.

$$\mathbf{b} = (\mathbf{J}^{\mathrm{T}}\mathbf{\Gamma}'\mathbf{J} + \mathbf{\Gamma}'')^{1/2} \tag{C.3}$$

However, finding the value of $\mathbf{c}$ is not as straighforward as before. By noting that each term of the identity given in equation C.2 is a scalar, further simplifications can be done.

$$\begin{aligned} \mathbf{K}^{\mathrm{T}}\mathbf{\Gamma}'\mathbf{J}\mathbf{Q}'' &= \big(\mathbf{K}^{\mathrm{T}}\mathbf{\Gamma}'\mathbf{J}\mathbf{Q}''\big)^{\mathrm{T}} = \mathbf{Q}''^{\mathrm{T}}\mathbf{J}^{\mathrm{T}}\mathbf{\Gamma}'\mathbf{K}' \\ \mathbf{c}^{\mathrm{T}}\mathbf{b}\mathbf{Q}'' &= \big(\mathbf{c}^{\mathrm{T}}\mathbf{b}\mathbf{Q}''\big)^{\mathrm{T}} = \mathbf{Q}''^{\mathrm{T}}\mathbf{b}^{\mathrm{T}}\mathbf{c} \end{aligned}$$

Using these both identities, it is now possible to easily equate the coefficients of $\mathbf{Q}''$:

$$\begin{aligned} 2\mathbf{c}^{\mathrm{T}}\mathbf{b} &= 2\mathbf{K}^{\mathrm{T}}\mathbf{\Gamma}'\mathbf{J} - 4\mathbf{U}^{\mathrm{T}}\mathbf{\Gamma}''^{1/2} - 4\mathbf{T}^{\mathrm{T}}\mathbf{\Gamma}'^{1/2}\mathbf{J} \\ \Rightarrow \mathbf{c}^{\mathrm{T}} &= \big(\mathbf{K}^{\mathrm{T}}\mathbf{\Gamma}'\mathbf{J} - 2\mathbf{U}^{\mathrm{T}}\mathbf{\Gamma}''^{1/2} - 2\mathbf{T}^{\mathrm{T}}\mathbf{\Gamma}'^{1/2}\mathbf{J}\big)(\mathbf{J}^{\mathrm{T}}\mathbf{\Gamma}'\mathbf{J} + \mathbf{\Gamma}'')^{-1/2} \end{aligned} \tag{C.4}$$

Finally, $\mathbf{d}$ can be formulated from the relation $\mathbf{c}^{\mathrm{T}}\mathbf{c} + \mathbf{d} = 0$

$$\mathbf{d} = -\big(\mathbf{K}^{\mathrm{T}}\mathbf{\Gamma}'\mathbf{J} - 2\mathbf{U}^{\mathrm{T}}\mathbf{\Gamma}''^{1/2} - 2\mathbf{T}^{\mathrm{T}}\mathbf{\Gamma}'^{1/2}\mathbf{J}\big)\big(\mathbf{J}^{\mathrm{T}}\mathbf{\Gamma}'\mathbf{J} + \mathbf{\Gamma}''\big)^{-1}\big(\mathbf{K}^{\mathrm{T}}\mathbf{\Gamma}'\mathbf{J} - 2\mathbf{U}^{\mathrm{T}}\mathbf{\Gamma}''^{1/2} - 2\mathbf{T}^{\mathrm{T}}\mathbf{\Gamma}'^{1/2}\mathbf{J}\big)^{\mathrm{T}} \tag{C.5}$$

For now, we will not develop further $\mathbf{d}$. By operating the change of variable and following the procedure given in appendix B, the power series given in equation C.1 can be written:

$$\begin{aligned} f(\mathbf{T},\mathbf{U}) = 2^{N/2}\det\big[\boldsymbol{\Gamma}'\,\boldsymbol{\Gamma}''\big]^{1/4}\left[\frac{\det(\mathbf{J})}{\det\big(\mathbf{J}^{\mathrm{T}}\boldsymbol{\Gamma}'\mathbf{J}+\boldsymbol{\Gamma}''\big)}\right]^{1/2}\exp\Big\{-\mathbf{U}^2-\mathbf{T}^2-\frac{1}{2}\mathbf{K}^{\mathrm{T}}\boldsymbol{\Gamma}'\mathbf{K} \\ +2\mathbf{T}^{\mathrm{T}}\boldsymbol{\Gamma}'^{1/2}\mathbf{K}+\frac{1}{2}\big[(\mathbf{K}^{\mathrm{T}}\boldsymbol{\Gamma}'\mathbf{J}-2\mathbf{U}^{\mathrm{T}}\boldsymbol{\Gamma}''^{1/2}-2\mathbf{T}^{\mathrm{T}}\boldsymbol{\Gamma}'^{1/2}\mathbf{J})(\mathbf{J}^{\mathrm{T}}\boldsymbol{\Gamma}'\mathbf{J}+\boldsymbol{\Gamma}'')^{-1} \\ \big(\mathbf{K}^{\mathrm{T}}\boldsymbol{\Gamma}'\mathbf{J}-2\mathbf{U}^{\mathrm{T}}\boldsymbol{\Gamma}''^{1/2}-2\mathbf{T}^{\mathrm{T}}\boldsymbol{\Gamma}'^{1/2}\mathbf{J}\big)^{\mathrm{T}}\big]\Big\} \end{aligned} \tag{C.6}$$

The terms in the exponential in equation C.6 can be reordered and grouped with respect to the dummy variables $\mathbf{U}$ and $\mathbf{T}$, reminding that each term of the exponiential is a scalar.

$$\begin{aligned} f(\mathbf{T},\mathbf{U}) = 2^{N/2}\det\big[\boldsymbol{\Gamma}'\,\boldsymbol{\Gamma}''\big]^{1/4}\left[\frac{\det(\mathbf{J})}{\det\big(\mathbf{J}^{\mathrm{T}}\boldsymbol{\Gamma}'\mathbf{J}+\boldsymbol{\Gamma}''\big)}\right]^{1/2} \\ \times\exp\Big[-\frac{1}{2}\mathbf{K}^{\mathrm{T}}\boldsymbol{\Gamma}'\mathbf{K}+\frac{1}{2}\mathbf{K}^{\mathrm{T}}\boldsymbol{\Gamma}'\mathbf{J}(\mathbf{J}^{\mathrm{T}}\boldsymbol{\Gamma}'\mathbf{J}+\boldsymbol{\Gamma}'')^{-1}\mathbf{J}^{\mathrm{T}}\boldsymbol{\Gamma}'\mathbf{K}\Big] \\ \times\exp\Big[\mathbf{U}^{\mathrm{T}}(2\boldsymbol{\Gamma}''^{1/2}(\mathbf{J}^{\mathrm{T}}\boldsymbol{\Gamma}'\mathbf{J}+\boldsymbol{\Gamma}'')^{-1}\boldsymbol{\Gamma}''^{1/2}-\mathbf{I})\mathbf{U}\Big] \\ \times\exp\Big[-2\mathbf{K}^{\mathrm{T}}\boldsymbol{\Gamma}''^{1/2}\mathbf{J}(\mathbf{J}^{\mathrm{T}}\boldsymbol{\Gamma}'\mathbf{J}+\boldsymbol{\Gamma}'')^{-1}\boldsymbol{\Gamma}''^{1/2}\mathbf{U}\Big] \\ \times\exp\Big[\mathbf{T}^{\mathrm{T}}(2\boldsymbol{\Gamma}'^{1/2}\mathbf{J}(\mathbf{J}^{\mathrm{T}}\boldsymbol{\Gamma}'\mathbf{J}+\boldsymbol{\Gamma}'')^{-1}\mathbf{J}^{\mathrm{T}}\boldsymbol{\Gamma}'^{1/2}-\mathbf{I})\mathbf{T}\Big] \\ \times\exp\Big[-2\mathbf{K}^{\mathrm{T}}(\boldsymbol{\Gamma}'\mathbf{J}(\mathbf{J}^{\mathrm{T}}\boldsymbol{\Gamma}'\mathbf{J}+\boldsymbol{\Gamma}'')^{-1}\mathbf{J}^{\mathrm{T}}-\mathbf{I})\boldsymbol{\Gamma}'^{1/2}\mathbf{T}\Big] \\ \times\exp\Big[\mathbf{U}^{\mathrm{T}}4\boldsymbol{\Gamma}''^{1/2}(\mathbf{J}^{\mathrm{T}}\boldsymbol{\Gamma}'\mathbf{J}+\boldsymbol{\Gamma}'')^{-1}\mathbf{J}^{\mathrm{T}}\boldsymbol{\Gamma}'^{1/2}\mathbf{T}\Big] \end{aligned} \tag{C.7}$$

where $\mathbf{I}$ is the identity matrix.

The first two lines of the right-hand side is the Franck-Condon integral of the vibrational ground states as given in equation 1.76. The five last exponentials are widely used for the generation of the analytic formulae in the method of Sharp and Rosenstock, so we will define 3 matrices ($\mathbf{A}$, $\mathbf{C}$, $\mathbf{E}$) and 2 vectors ($\mathbf{B}$, $\mathbf{D}$) as:

$$\mathbf{A} = 2\boldsymbol{\Gamma}'^{1/2}\mathbf{J}(\mathbf{J}^{\mathrm{T}}\boldsymbol{\Gamma}'\mathbf{J}+\boldsymbol{\Gamma}'')^{-1}\mathbf{J}^{\mathrm{T}}\boldsymbol{\Gamma}'^{1/2}-\mathbf{I} \tag{C.8}$$

$$\mathbf{B} = -2\boldsymbol{\Gamma}'^{1/2}(\mathbf{J}(\mathbf{J}^{\mathrm{T}}\boldsymbol{\Gamma}'\mathbf{J}+\boldsymbol{\Gamma}'')^{-1}\mathbf{J}^{\mathrm{T}}\boldsymbol{\Gamma}'-\mathbf{I})\mathbf{K} \tag{C.9}$$

$$\mathbf{C} = 2\boldsymbol{\Gamma}''^{1/2}(\mathbf{J}^{\mathrm{T}}\boldsymbol{\Gamma}'\mathbf{J}+\boldsymbol{\Gamma}'')^{-1}\boldsymbol{\Gamma}''^{1/2}-\mathbf{I} \tag{C.10}$$

$$\mathbf{D} = -2\boldsymbol{\Gamma}''^{1/2}(\mathbf{J}^{\mathrm{T}}\boldsymbol{\Gamma}'\mathbf{J}+\boldsymbol{\Gamma}'')^{-1}\mathbf{J}^{\mathrm{T}}\boldsymbol{\Gamma}''^{1/2}\mathbf{K} \tag{C.11}$$

$$\mathbf{E} = 4\boldsymbol{\Gamma}''^{1/2}(\mathbf{J}^{\mathrm{T}}\boldsymbol{\Gamma}'\mathbf{J}+\boldsymbol{\Gamma}'')^{-1}\mathbf{J}^{\mathrm{T}}\boldsymbol{\Gamma}'^{1/2} \tag{C.12}$$

Using these variables, equation C.7 can be simply written:

$$f(\mathbf{T},\mathbf{U}) = \langle\,\mathbf{0}'\mid\mathbf{0}''\,\rangle\exp\Big[\mathbf{U}^{\mathrm{T}}\mathbf{C}\mathbf{U}+\mathbf{D}^{\mathrm{T}}\mathbf{U}+\mathbf{T}^{\mathrm{T}}\mathbf{A}\mathbf{T}+\mathbf{B}^{\mathrm{T}}\mathbf{T}+\mathbf{U}^{\mathrm{T}}\mathbf{E}\mathbf{T}\Big] \tag{C.13}$$

A useful property of matrix $\mathbf{A}$ and $\mathbf{C}$ is their symmetry, while $\mathbf{E}$ is nonsymmetric. It is straightforwardly shown by evaluating the transpose of these matrices, recalling that $\mathbf{b}$ is symmetric.

Appendix D

Calculation of the transition dipole moment integral in the case of D2

In this appendix, we will consider the calculation of the transition dipole moment integrals when the electronic transition dipole moment is described by a Taylor series up to the second order as shown in equation 1.29. Our calculations will be restricted to the case of the diagonal elements available in GAUSSIAN.

To make easier the reading of the following discussions, we will adopt some unformal notations similar to the one used in section 3.7:

$$\dot{\boldsymbol{\mu}}_k = \left(\frac{\partial \boldsymbol{\mu}_{if}}{\partial Q''_k}\right)_0 \sqrt{\frac{\hbar}{2\omega''_k}}$$
$$\ddot{\boldsymbol{\mu}}_{kk} = \left(\frac{\partial^2 \boldsymbol{\mu}_{if}}{\partial Q''^2_k}\right)_0 \frac{\hbar}{2\omega''_k}$$

First, we will restrict our case to the second derivative of the electronic transition dipole moment (*type of calculation*: `D2`):

$$\begin{aligned}\langle\, \boldsymbol{v}' \mid \boldsymbol{\mu}_{if} \mid \boldsymbol{v}'' \,\rangle = \sum_{k=1}^{N} \frac{\ddot{\boldsymbol{\mu}}_{kk}}{2} \Big[&\sqrt{v''_k(v''_k-1)}\, \langle\, \boldsymbol{v}' \mid \boldsymbol{v}'' - 2''_k \,\rangle \\ &+ (2v''_k+1)\, \langle\, \boldsymbol{v}' \mid \boldsymbol{v}'' \,\rangle + \sqrt{(v''_k+1)(v''_k+2)}\, \langle\, \boldsymbol{v}' \mid \boldsymbol{v}'' + 2''_k \,\rangle \Big]\end{aligned} \tag{D.1}$$

In reccurence treatment, $\langle\, \boldsymbol{v}' \mid \boldsymbol{v}'' + 2''_k \,\rangle$ must be expressed with respect to lower quanta. This is done using the recursion formula given in equation 2.31:

$$\begin{aligned}\langle\, \boldsymbol{v}' \mid \boldsymbol{v}'' + 2''_k \,\rangle = \frac{1}{\sqrt{2(v''_k+2)}} \Big[&\mathrm{D}_k \langle\, \boldsymbol{v}' \mid \boldsymbol{v}'' + 1''_k \,\rangle + \sqrt{2(v''_k+1)}\, \mathrm{C}_{kk} \langle\, \boldsymbol{v}' \mid \boldsymbol{v}'' \,\rangle \\ &+ \sum_{\substack{l=1 \\ l\neq k}}^{N} \sqrt{2v''_l}\, \mathrm{C}_{kl} \langle\, \boldsymbol{v}' \mid \boldsymbol{v}'' + 1''_k - 1''_l \,\rangle + \sum_{l=1}^{N} \sqrt{v'_l}\, \mathrm{E}_{kl} \langle\, \boldsymbol{v}' - 1'_l \mid \boldsymbol{v}'' + 1''_k \,\rangle \Big]\end{aligned} \tag{D.2}$$

We need to apply reccurence relations again to express three kinds of overlap integrals with

respect to lower quanta:

$$\begin{aligned}\langle\, \boldsymbol{v}' \mid \boldsymbol{v}'' + 1''_k \,\rangle = \frac{1}{\sqrt{2(v''_k+1)}}\Big[&\mathrm{D}_k\langle\, \boldsymbol{v}' \mid \boldsymbol{v}'' \,\rangle + \sum_{l=1}^{N}\sqrt{2v''_l}\;\mathrm{C}_{kl}\langle\, \boldsymbol{v}' \mid \boldsymbol{v}'' - 1''_l \,\rangle \\ &+ \sum_{l=1}^{N}\sqrt{v'_l}\;\mathrm{E}_{kl}\langle\, \boldsymbol{v}' - 1'_l \mid \boldsymbol{v}'' \,\rangle\Big]\end{aligned} \tag{D.3}$$

$$\begin{aligned}\langle\, \boldsymbol{v}' \mid \boldsymbol{v}'' + 1''_k - 1''_l \,\rangle = \frac{1}{\sqrt{2(v''_k+1)}}\Big[&\mathrm{D}_k\langle\, \boldsymbol{v}' \mid \boldsymbol{v}'' - 1''_l \,\rangle + \sqrt{2(v''_l-1)}\;\mathrm{C}_{kl}\langle\, \boldsymbol{v}' \mid \boldsymbol{v}'' - 2''_l \,\rangle \\ &+ \sum_{\substack{m=1\\m\neq l}}^{N}\sqrt{2v''_m}\;\mathrm{C}_{km}\langle\, \boldsymbol{v}' \mid \boldsymbol{v}'' - 1''_l - 1''_m \,\rangle + \sum_{l=1}^{N}\sqrt{v'_m}\;\mathrm{E}_{km}\langle\, \boldsymbol{v}' - 1'_m \mid \boldsymbol{v}'' - 1''_l \,\rangle\Big]\end{aligned} \tag{D.4}$$

$$\begin{aligned}\langle\, \boldsymbol{v}' - 1'_l \mid \boldsymbol{v}'' + 1''_k \,\rangle = \frac{1}{\sqrt{2(v''_k+1)}}\Big[&\mathrm{D}_k\langle\, \boldsymbol{v}' - 1'_l \mid \boldsymbol{v}'' \,\rangle + \sum_{m=1}^{N}\sqrt{2v''_m}\;\mathrm{C}_{km}\langle\, \boldsymbol{v}' - 1'_l \mid \boldsymbol{v}'' - 1''_m \,\rangle \\ &+ \sqrt{v'_l-1}\;\mathrm{E}_{kl}\langle\, \boldsymbol{v}' - 2'_l \mid \boldsymbol{v}'' \,\rangle + \sum_{\substack{m=1\\m\neq l}}^{N}\sqrt{v'_m}\;\mathrm{E}_{km}\langle\, \boldsymbol{v}' - 1'_l - 1'_m \mid \boldsymbol{v}'' \,\rangle\Big]\end{aligned} \tag{D.5}$$

Restricting our study to the neglect of temperature, as it is currently the case in our procedure, then the last term in the right-hand side of equations D.2, D.3 and D.4 are cancelled and equation D.5 is not taken into account. Introducing equation D.2 and D.3 in equation D.2, one obtains:

$$\begin{aligned}\langle\, \mathbf{0}' \mid \boldsymbol{v}'' + 2''_k \,\rangle = \frac{1}{\sqrt{2(v''_k+2)}}\Bigg\{&\mathrm{D}_k \times \Bigg(\frac{1}{\sqrt{2(v''_k+1)}}\Big[\mathrm{D}_k\langle\, \mathbf{0}' \mid \boldsymbol{v}'' \,\rangle \\ &+ \sum_{l=1}^{N}\sqrt{2v''_l}\;\mathrm{C}_{kl}\langle\, \mathbf{0}' \mid \boldsymbol{v}'' - 1''_l \,\rangle\Big]\Bigg) \\ &+ \sqrt{2(v''_k+1)}\;\mathrm{C}_{kk}\langle\, \mathbf{0}' \mid \boldsymbol{v}'' \,\rangle \\ &+ \sum_{\substack{l=1\\l\neq k}}^{N}\sqrt{2v''_l}\;\mathrm{C}_{kl} \times \Bigg(\frac{1}{\sqrt{2(v''_k+1)}}\Big[\mathrm{D}_k\langle\, \mathbf{0}' \mid \boldsymbol{v}'' - 1''_l \,\rangle \\ &+ \sqrt{2(v''_l-1)}\;\mathrm{C}_{kl}\langle\, \mathbf{0}' \mid \boldsymbol{v}'' - 2''_l \,\rangle \\ &+ \sum_{\substack{m=1\\m\neq l}}^{N}\sqrt{2v''_m}\;\mathrm{C}_{km}\langle\, \mathbf{0}' \mid \boldsymbol{v}'' - 1''_l - 1''_m \,\rangle\Big]\Bigg)\Bigg\}\end{aligned} \tag{D.6}$$

Using equation D.6, a fully usable recurrence formula can be expressed in place of equation D.1.

Upon reordering the terms with the same overlap integrals, one obtains:

$$\begin{aligned}
\langle\, \mathbf{0}' \mid \boldsymbol{\mu}_{if} \mid \boldsymbol{v}'' \,\rangle = \sum_{k=1}^{N} \frac{\ddot{\boldsymbol{\mu}}_{kk}}{2} \Bigg[& \sqrt{v_k''(v_k''-1)}\, \langle\, \mathbf{0}' \mid \boldsymbol{v}'' - 2_k'' \,\rangle \\
& + \sum_{\substack{l=1 \\ l\neq k}}^{N} \sqrt{v_l''(v_l''-1)}\, \mathrm{C}_{kl}{}^2 \langle\, \mathbf{0}' \mid \boldsymbol{v}'' - 2_l'' \,\rangle \\
& + \sum_{\substack{l=1 \\ l\neq k}}^{N} \sum_{\substack{m=1 \\ m\neq l}}^{N} \sqrt{v_l'' v_m''}\, \mathrm{C}_{kl}\mathrm{C}_{km} \langle\, \mathbf{0}' \mid \boldsymbol{v}'' - 1_l'' - 1_m'' \,\rangle \\
& + \sum_{l=1}^{N} \sqrt{2v_l''}\mathrm{C}_{kl}\mathrm{D}_k \langle\, \mathbf{0}' \mid \boldsymbol{v}'' - 1_l'' \,\rangle \\
& + \left(2v_k'' + 1 + \frac{\mathrm{D}_k}{2} + (v_k''+1) \times \mathrm{C}_{kk}\right) \langle\, \mathbf{0}' \mid \boldsymbol{v}'' \,\rangle \Bigg]
\end{aligned} \tag{D.7}$$

In our computational strategy, the transition dipole moment integrals are calculated inside a *set* of some class $\mathcal{C}_n$. In this framework, equation D.7 is given by the relation:

$$\begin{aligned}
\langle\, \mathbf{0}' \mid \boldsymbol{\mu}_{if} \mid \boldsymbol{v}'' \,\rangle = \sum_{k=1}^{N} \frac{\ddot{\boldsymbol{\mu}}_{kk}}{2} \Bigg[& \sqrt{v_k''(v_k''-1)}\, \langle\, \mathbf{0}' \mid \boldsymbol{v}'' - 2_k'' \,\rangle \\
& + \sum_{\substack{l=1 \\ \underline{v}_{l_{\mathrm{id}}}'' \neq k}}^{n} \sqrt{\underline{v}_l''(\underline{v}_l''-1)}\, \mathrm{C}(k, \underline{v}_{l_{\mathrm{id}}}'')^2 \langle\, \mathbf{0}' \mid \underline{\boldsymbol{v}}'' - 2_l'' \,\rangle \\
& + \sum_{\substack{l=1 \\ \underline{v}_{l_{\mathrm{id}}}'' \neq k}}^{n} \sum_{\substack{m=1 \\ m\neq l}}^{n} \sqrt{\underline{v}_l'' \underline{v}_m''}\, \mathrm{C}(k, \underline{v}_{l_{\mathrm{id}}})\mathrm{C}(k, \underline{v}_{m_{\mathrm{id}}}'') \langle\, \mathbf{0}' \mid \underline{\boldsymbol{v}}'' - 1_l'' - 1_m'' \,\rangle \\
& + \sum_{l=1}^{n} \sqrt{2\underline{v}_l''}\, \mathrm{C}(k, \underline{v}_{l_{\mathrm{id}}}'')\mathrm{D}_k \langle\, \mathbf{0}' \mid \underline{\boldsymbol{v}}'' - 1_l'' \,\rangle \\
& + \left(2v_k'' + 1 + \frac{\mathrm{D}_k}{2} + (v_k''+1) \times \mathrm{C}_{kk}\right) \langle\, \mathbf{0}' \mid \underline{\boldsymbol{v}}'' \,\rangle \Bigg]
\end{aligned} \tag{D.8}$$

Let us now consider the general case of any *type of calculation* up to D2. To do so, we add equation D.8 to the formula of the transition dipole moment integral given for the case of FCHT

calculations in section 3.7. However, instead of equation 3.15, we will use the clearer equation 3.14:

$$
\begin{aligned}
\langle\, \mathbf{0}' \mid \boldsymbol{\mu}_{if} \mid \boldsymbol{v}'' \,\rangle = & \left[\boldsymbol{\mu}_{if}(\mathbf{Q}_0'') + \sum_{k=1}^{N} \dot{\boldsymbol{\mu}}_k \frac{\mathrm{D}_k}{\sqrt{2}}\right] \langle\, \mathbf{0}' \mid \underline{\boldsymbol{v}}'' \,\rangle \\
& + \sum_{k=1}^{n} \left[\underline{\dot{\boldsymbol{\mu}}}_k + \sum_{l=1}^{N} \dot{\boldsymbol{\mu}}_l \mathrm{C}(l, \underline{v}''_{k_{\mathrm{id}}})\right] \sqrt{\underline{v}''_k}\, \langle\, \mathbf{0}' \mid \underline{\boldsymbol{v}}'' - 1''_k \,\rangle \\
& + \sum_{k=1}^{N} \frac{\ddot{\boldsymbol{\mu}}_{kk}}{2} \left[\sqrt{v''_k(v''_k - 1)}\, \langle\, \mathbf{0}' \mid \boldsymbol{v}'' - 2''_k \,\rangle \right. \\
& \quad + \sum_{\substack{l=1 \\ \underline{v}''_{l_{\mathrm{id}}} \neq k}}^{n} \sqrt{\underline{v}''_l(\underline{v}''_l - 1)}\, \mathrm{C}(k, \underline{v}''_{l_{\mathrm{id}}})^2 \langle\, \mathbf{0}' \mid \underline{\boldsymbol{v}}'' - 2''_l \,\rangle \\
& \quad + \sum_{\substack{l=1 \\ \underline{v}''_{l_{\mathrm{id}}} \neq k}}^{n} \sum_{\substack{m=1 \\ m \neq l}}^{n} \sqrt{\underline{v}''_l \underline{v}''_m}\, \mathrm{C}(k, \underline{v}_{l_{\mathrm{id}}}) \mathrm{C}(k, \underline{v}''_{m_{\mathrm{id}}}) \langle\, \mathbf{0}' \mid \underline{\boldsymbol{v}}'' - 1''_l - 1''_m \,\rangle \\
& \quad + \sum_{l=1}^{n} \sqrt{2\underline{v}''_l}\, \mathrm{C}(k, \underline{v}''_{l_{\mathrm{id}}}) \mathrm{D}_k \langle\, \mathbf{0}' \mid \underline{\boldsymbol{v}}'' - 1''_l \,\rangle \\
& \quad \left. + \left(2v''_k + 1 + \frac{\mathrm{D}_k}{2} + (v''_k + 1) \times \mathrm{C}_{kk}\right) \langle\, \mathbf{0}' \mid \underline{\boldsymbol{v}}'' \,\rangle \right]
\end{aligned}
\tag{D.9}
$$

By interchanging indexes in some terms, a first shortening of the formula above can be done:

$$
\begin{aligned}
\langle\, \mathbf{0}' \mid \boldsymbol{\mu}_{if} \mid \boldsymbol{v}'' \,\rangle = & \left[\boldsymbol{\mu}_{if}(\mathbf{Q}_0'') + \sum_{k=1}^{N} \dot{\boldsymbol{\mu}}_k \frac{\mathrm{D}_k}{\sqrt{2}} \right. \\
& \quad \left. + \sum_{k=1}^{N} \frac{\ddot{\boldsymbol{\mu}}_{kk}}{2} \left(2v''_k + 1 + \frac{\mathrm{D}_k}{2} + (v''_k + 1) \times \mathrm{C}_{kk}\right)\right] \langle\, \mathbf{0}' \mid \underline{\boldsymbol{v}}'' \,\rangle \\
& + \sum_{k=1}^{n} \left[\underline{\dot{\boldsymbol{\mu}}}_k + \sum_{l=1}^{N} \left\{\dot{\boldsymbol{\mu}}_l \mathrm{C}(l, \underline{v}''_{k_{\mathrm{id}}}) + \frac{\ddot{\boldsymbol{\mu}}_{ll}}{2} \sqrt{2}\, \mathrm{C}(l, \underline{v}''_{k_{\mathrm{id}}}) \mathrm{D}_l \right\}\right] \sqrt{\underline{v}''_k}\, \langle\, \mathbf{0}' \mid \underline{\boldsymbol{v}}'' - 1''_k \,\rangle \\
& + \sum_{k=1}^{N} \frac{\ddot{\boldsymbol{\mu}}_{kk}}{2} \left[\sqrt{v''_k(v''_k - 1)}\, \langle\, \mathbf{0}' \mid \boldsymbol{v}'' - 2''_k \,\rangle \right. \\
& \quad + \sum_{\substack{l=1 \\ \underline{v}''_{l_{\mathrm{id}}} \neq k}}^{n} \sqrt{\underline{v}''_l(\underline{v}''_l - 1)}\, \mathrm{C}(k, \underline{v}''_{l_{\mathrm{id}}})^2 \langle\, \mathbf{0}' \mid \underline{\boldsymbol{v}}'' - 2''_l \,\rangle \\
& \quad \left. + \sum_{\substack{l=1 \\ \underline{v}''_{l_{\mathrm{id}}} \neq k}}^{n} \sum_{\substack{m=1 \\ m \neq l}}^{n} \sqrt{\underline{v}''_l \underline{v}''_m}\, \mathrm{C}(k, \underline{v}_{l_{\mathrm{id}}}) \mathrm{C}(k, \underline{v}''_{m_{\mathrm{id}}}) \langle\, \mathbf{0}' \mid \underline{\boldsymbol{v}}'' - 1''_l - 1''_m \,\rangle \right]
\end{aligned}
\tag{D.10}
$$

The last three terms in the right-hand side of equation D.10 are more complex to handle because of the partial summations. It is noteworthy that the condition $\underline{v}''_{l_{\mathrm{id}}} = k$ can only be met if mode k belongs to the *set*, otherwise it can be safely disregarded. Consequently, for the first summation on the N modes, two cases exist: $k \in Z_n$ or $k \notin Z_n$ where Z_n is a group containing all the mode

selected in a *set* of dimension n. For $k \notin Z_n$, we can write the following relations:

$$\begin{aligned}
&\sum_{k\notin Z_n} \frac{\ddot{\boldsymbol{\mu}}_{kk}}{2} \sum_{l=1}^{n} \sqrt{\underline{v}''_l(\underline{v}''_l-1)}\,\mathrm{C}(k,\underline{v}''_{l_{\mathrm{id}}})^2 \langle\, \mathbf{0'} \mid \underline{\boldsymbol{v}}'' - 2''_l \,\rangle \\
&\qquad = \sum_{k=1}^{n} \sqrt{\underline{v}''_k(\underline{v}''_k-1)} \sum_{l\notin Z_n} \frac{\ddot{\boldsymbol{\mu}}_{ll}}{2}\,\mathrm{C}(l,\underline{v}''_{k_{\mathrm{id}}})^2 \langle\, \mathbf{0'} \mid \underline{\boldsymbol{v}}'' - 2''_k \,\rangle \\
&\sum_{k\notin Z_n} \frac{\ddot{\boldsymbol{\mu}}_{kk}}{2} \sum_{l=1}^{n} \sum_{\substack{m=1\\m\neq l}}^{n} \sqrt{\underline{v}''_l\underline{v}''_m}\,\mathrm{C}(k,\underline{v}_{l_{\mathrm{id}}})\mathrm{C}(k,\underline{v}''_{m_{\mathrm{id}}}) \langle\, \mathbf{0'} \mid \underline{\boldsymbol{v}}'' - 1''_l - 1''_m \,\rangle \\
&\qquad = \sum_{k=1}^{n} \sum_{\substack{l=1\\l\neq k}}^{n} \sqrt{\underline{v}''_k\underline{v}''_l} \sum_{m\notin Z_n} \frac{\ddot{\boldsymbol{\mu}}_{mm}}{2}\,\mathrm{C}(m,\underline{v}_{k_{\mathrm{id}}})\mathrm{C}(m,\underline{v}''_{l_{\mathrm{id}}}) \langle\, \mathbf{0'} \mid \underline{\boldsymbol{v}}'' - 1''_k - 1''_l \,\rangle
\end{aligned}$$

if $k \in Z_n$, the conditions on the summation cannot be overcome simply and can be let in such a way. Finally, it is possible to write equation D.10 as:

$$\begin{aligned}
\langle\, \mathbf{0'} \mid \boldsymbol{\mu}_{if} \mid \boldsymbol{v}'' \,\rangle = &\left[\boldsymbol{\mu}_{if}(\mathbf{Q}''_0) + \sum_{k=1}^{N} \dot{\boldsymbol{\mu}}_k \frac{\mathrm{D}_k}{\sqrt{2}} \right.\\
&\left. + \sum_{k=1}^{N} \frac{\ddot{\boldsymbol{\mu}}_{kk}}{2}\left(2v''_k + 1 + \frac{\mathrm{D}_k}{2} + (v''_k+1)\times \mathrm{C}_{kk}\right)\right] \langle\, \mathbf{0'} \mid \underline{\boldsymbol{v}}'' \,\rangle \\
&+ \sum_{k=1}^{n}\left[\underline{\dot{\boldsymbol{\mu}}}_k + \sum_{l=1}^{N}\left\{\dot{\boldsymbol{\mu}}_l \mathrm{C}(l,\underline{v}''_{k_{\mathrm{id}}}) + \frac{\ddot{\boldsymbol{\mu}}_{ll}}{2}\sqrt{2}\,\mathrm{C}(l,\underline{v}''_{k_{\mathrm{id}}})\mathrm{D}_l\right\}\right]\sqrt{\underline{v}''_k}\,\langle\, \mathbf{0'} \mid \underline{\boldsymbol{v}}'' - 1''_k \,\rangle \\
&+ \sum_{k=1}^{n}\left[1 + \sum_{l\notin Z_n} \frac{\ddot{\boldsymbol{\mu}}_{ll}}{2}\,\mathrm{C}(l,\underline{v}''_{k_{\mathrm{id}}})^2\right]\sqrt{\underline{v}''_k(\underline{v}''_k-1)}\langle\, \mathbf{0'} \mid \underline{\boldsymbol{v}}'' - 2''_k \,\rangle \\
&+ \sum_{k=1}^{n}\sum_{\substack{l=1\\l\neq k}}^{n} \sqrt{\underline{v}''_k\underline{v}''_l} \sum_{m\notin Z_n} \frac{\ddot{\boldsymbol{\mu}}_{mm}}{2}\,\mathrm{C}(m,\underline{v}_{k_{\mathrm{id}}})\mathrm{C}(m,\underline{v}''_{l_{\mathrm{id}}}) \langle\, \mathbf{0'} \mid \underline{\boldsymbol{v}}'' - 1''_k - 1''_l \,\rangle \\
&+ \sum_{k\in Z_n}^{N} \frac{\ddot{\boldsymbol{\mu}}_{kk}}{2} \sum_{\substack{l=1\\\underline{v}''_{l_{\mathrm{id}}}\neq k}}^{n} \sqrt{\underline{v}''_l(\underline{v}''_l-1)}\,\mathrm{C}(k,\underline{v}''_{l_{\mathrm{id}}})^2 \langle\, \mathbf{0'} \mid \underline{\boldsymbol{v}}'' - 2''_l \,\rangle \\
&+ \sum_{k\in Z_n}^{N} \frac{\ddot{\boldsymbol{\mu}}_{kk}}{2} \sum_{\substack{l=1\\\underline{v}''_{l_{\mathrm{id}}}\neq k}}^{n} \sum_{\substack{m=1\\m\neq l}}^{n} \sqrt{\underline{v}''_l\underline{v}''_m}\,\mathrm{C}(k,\underline{v}_{l_{\mathrm{id}}})\mathrm{C}(k,\underline{v}''_{m_{\mathrm{id}}}) \langle\, \mathbf{0'} \mid \underline{\boldsymbol{v}}'' - 1''_l - 1''_m \,\rangle
\end{aligned} \tag{D.11}$$

Some computational strategies can be devised to speed up the calculations. The first one is to compute only once the factor of $\langle\, \mathbf{0'} \mid \boldsymbol{v}'' \,\rangle$:

$$\boldsymbol{\mu}_{\mathrm{cst}} = \boldsymbol{\mu}_{if}(\mathbf{Q}''_0) + \sum_{k=1}^{N} \dot{\boldsymbol{\mu}}_k \frac{\mathrm{D}(\underline{v}''_{k_{\mathrm{id}}})}{\sqrt{2}} + \sum_{k=1}^{N} \frac{\ddot{\boldsymbol{\mu}}_{kk}}{2}\left(2v''_k + 1 + \frac{\mathrm{D}_k}{2} + (v''_k+1)\times \mathrm{C}_{kk}\right)$$

A seeming issue for the speed lies on the distinction between $k \in Z_n$ and $k \notin Z_n$ for the sums. A straightforward solution is to extend the dimension of $\underline{\boldsymbol{v}}''$ from n to N. The first n elements correspond to the modes treated in the *set* and the $N-n$ followine elements are the remaining modes. The sorting is done by our loop described at the end of section 3.7. As a consequence, $\underline{\boldsymbol{v}}''$

is larger than previously. However, as we will see below, it is unnecessary to increase the dimension of $\underline{\boldsymbol{v}}''_{\mathrm{id}}$.

Finally, to take full advantage of the small matrices, it is better to create two new matrices defined as:

$$\underline{\underline{\mathrm{C}}}_{kl} = \sqrt{\frac{\ddot{\boldsymbol{\mu}}_{kk}}{2}}\mathrm{C}(\underline{v}_{k_{\mathrm{id}}}, \underline{v}''_{l_{\mathrm{id}}})$$
$$\underline{\underline{\mathrm{D}}}_{k} = \sqrt{\frac{\ddot{\boldsymbol{\mu}}_{kk}}{2}}\mathrm{D}(\underline{v}_{k_{\mathrm{id}}})$$

Constructing these matrices, we do not need to know the equivalence of the indexes in $\underline{\boldsymbol{v}}''$ and $\boldsymbol{v}''$, which means that the vector $\underline{\boldsymbol{v}}''_{\mathrm{id}}$ is obsolete. Taking into account the strategies described above and in section 3.7, the transition dipole moment integral is given by the formula:

$$\begin{aligned}
\langle\, \mathbf{0}' \mid \boldsymbol{\mu}_{if} \mid \boldsymbol{v}'' \,\rangle &= \boldsymbol{\mu}_{\mathrm{cst}} \,\langle\, \mathbf{0}' \mid \underline{\boldsymbol{v}}'' \,\rangle \\
&+ \sum_{k=1}^{n}\left[\underline{\dot{\boldsymbol{\mu}}}_k + \sum_{l=1}^{N}\left\{\underline{\underline{\mathrm{C}}}_{lk} + \sqrt{2}\,\underline{\underline{\mathrm{C}}}_{lk}\underline{\underline{\mathrm{D}}}_{l}\right]\sqrt{\underline{v}''_k}\,\langle\, \mathbf{0}' \mid \underline{\boldsymbol{v}}'' - 1''_k \,\rangle \right. \\
&+ \sum_{k=1}^{n}\left[1 + \sum_{l=n+1}^{N}\frac{\ddot{\boldsymbol{\mu}}_{ll}}{2}\,\underline{\underline{\mathrm{C}}}_{lk}^{\,2}\right]\sqrt{\underline{v}''_k(\underline{v}''_k - 1)}\langle\, \mathbf{0}' \mid \underline{\boldsymbol{v}}'' - 2''_k \,\rangle \\
&+ \sum_{k=1}^{n}\sum_{\substack{l=1\\l\neq k}}^{n}\sqrt{\underline{v}''_k\underline{v}''_l}\sum_{m=n+1}^{N}\underline{\underline{\mathrm{C}}}_{mk}\underline{\underline{\mathrm{C}}}_{ml}\langle\, \mathbf{0}' \mid \underline{\boldsymbol{v}}'' - 1''_k - 1''_l \,\rangle \qquad \text{(D.12)} \\
&+ \sum_{k=1}^{n}\sum_{\substack{l=1\\l\neq k}}^{n}\sqrt{\underline{v}''_l(\underline{v}''_l - 1)}\,\underline{\underline{\mathrm{C}}}_{kl}^{\,2} \\
&+ \sum_{k=1}^{n}\sum_{\substack{l=1\\l\neq k}}^{n}\sum_{\substack{m=1\\m\neq l}}^{n}\sqrt{\underline{v}''_l\underline{v}''_m}\,\underline{\underline{\mathrm{C}}}_{kl}\underline{\underline{\mathrm{C}}}_{km}\langle\, \mathbf{0}' \mid \underline{\boldsymbol{v}}'' - 1''_l - 1''_m \,\rangle
\end{aligned}$$

The additional memory consumption can be quite cumbersome at this level. However, discarding the matrix $\underline{\underline{\mathbf{C}}}$ and the vector $\underline{\underline{\mathbf{D}}}$ implies that the subset $\underline{\boldsymbol{v}}''$ cannot be described as previously or $\underline{\boldsymbol{v}}''_{\mathrm{id}}$ must be expanded to contain all elements of $\underline{\boldsymbol{v}}''$.

Appendix E

List of the keywords currently recognized by FRANCK

In this section are summarized the keywords available in the procedure FRANCK along with their default value and the type of data expected in brackets. Except for the type of calculations, all keywords are optional.

Type of calculations

`FC/FCHT/HT` $\left[INTEGER, 1\right]$: Selects the terms of the Taylor expansion to approximate the transition dipole moment.

Type of transition

`ABS/ABSORPTION` $\left[INTEGER, 1\right]$: Sets the type of transition to be absorption

`EMI/EMISSION` $\left[INTEGER, 3\right]$: Sets the type of transition to be emission

Parameters for the electronic transition dipole moment (only available with READ2)

`NSTATE` $\left[INTEGER, 1\right]$: Sets which state should be read for the electronic transition dipole moment and its derivatives. This keyword is only taken into account if data for the final state are read from a Gaussian output file.

Parameters for the *Classes*[a] method

`MAXC1` $\left[INTEGER, 20\right]$: Sets the maximum number of quanta to reach when calculating the transitions corresponding to overtones ($\mathcal{C}_{1_{\max}}$).

[a] An *a priori* selection of transitions to compute based on an evaluation pool of overtones and two-states combinations

`MAXC2` $\left[INTEGER, 13\right]$: Sets the maximum quantum numbers to reach for both modes in combination bands of two simultaneously excited modes of the final state ($\mathcal{C}_{1_{\max}}$).

`MAXINT` $\left[INTEGER, 100\right]$: Sets the maximum number of transtions (in millions) computed for each *class* ($N_I^{\max}$).

`MAXBANDS` $\left[INTEGER, 7\right]$: Sets the highest class to compute

`SPDelta` $\left[FLOAT, 0.0\right]$: Forces the termination of the calculation if the relative spectrum convergence between two consecutive *classes* is lower that the given threshold ($\Delta_{SP} = SP_n - SP_{n-1}$). A value lower or equal to 0.0 deactivates the control.

`NOINTAN` $\left[BOOLEAN, FALSE\right]$: Deactivates the use of the Sharp and Rosenstock analytic formulae to compute overlap integrals in the case of single overtones and two-states combinations.

Parameters for the spectrum layout

`NORELI00` $\left[BOOLEAN, FALSE\right]$: By default, spectra bounds are given with respect to the energy of the 0-0 transition. This keyword must be given if absolute energies are given as spectrum bounds by the user.

`SPECMIN` $\left[FLOAT, -1000\right]$: Sets the lower bound (in cm^{-1}) of the final photoelectron spectrum.

`SPECMAX` $\left[FLOAT, +8000\right]$: Sets the upper bound (in cm^{-1}) of the final photoelectron spectrum.

`SPECRES` $\left[FLOAT, 8\right]$: Sets the gap (in cm^{-1}) between two points of the discretized spectrum. This value can greatly influence the times of computations, very low values slowing greatly the calculations, especially if `SPECHWHM` is set high.

`SPECHWHM` $\left[FLOAT, 135\right]$: Sets the Half-Width at Half-Maximum (in cm^{-1}) for the convolution of the spectral bands expressed with a Gaussian function.

Selection of the input source

`CALC1/CHK1/READ1` $\left[INTEGER, 2\right]$: Selects the source of data for the initial state between the current calculation, the GAUSSIAN checkpoint file or a GAUSSIAN 03/DV output file (given below). By default, the checkpoint file is used.

`CALC2/CHK2/READ2` $\left[INTEGER, 1\right]$: Selects the source of data for the final state between the current calculation, the GAUSSIAN checkpoint file or a GAUSSIAN

03/DV output file (given below). By default, the current calculation is used.

Parameters of the output printing

`ALLSPECTRA` $[BOOLEAN, FALSE]$: Prints in the GAUSSIAN output the resulting spectra for each set of combinations (*class*) in addition to the final spectrum. This printing is deactivated by default.

`PRTMAT` $[INTEGER, 0]$: A succession of figures to print different matrices used as a basis for integrals calculations : *1* for the Duschinsky matrix **J**, *2* for the shift vector **K**, *3* for **A**, *4* for **B**, *5* for **C**, *6* for **D** and *7* for **E**, where **A**, **B**, **C**, **D**, **E** are the Sharp and Rosenstock matrices. The order of the figures is not important.

`PRTINT` $[FLOAT, 0.01]$: Sets which integrals should be printed in output. The threshold is a fraction of the I_{00} intensity.

Interaction with the computational parameters

`InFrS0` $[BOOLEAN, FALSE]$: Forces the program to use frequencies given by the user for the initial state. These frequencies are given after the *dinautil section*.

`InFrS1` $[BOOLEAN, FALSE]$: Forces the program to use frequencies given by the user for the final state. These frequencies are given after the *dinautil section*.

`JDusch/JIdent` $[INTEGER, 0]$: Forces the program to use the normal Duschinsky matrix (`JDusch`) or an identity matrix as the Duschinsky matrix (`JIdent`). In the latter case, rotation of the normal modes is not taken into account.

`SclVec` $[BOOLEAN, FALSE]$: Enables computation of a scaling vector to modify frequencies of the final states using the scaling vector of the frequencies of the initial state and the Duschinsky matrix. When this keyword, is given, user frequencies are asked for the initial state in the same way as `InFrS0`.

`ENERINP` $[FLOAT, 0.]$: Replaces the computed ΔE between initial and final states by a user-given one.

Appendix F

Supplementary data for the study of the $^2B_1 \rightarrow {}^2A_2$ transition of the phenoxyl radical

F.1 GAUSSIAN Input files

State 1 (step3): Final state

```
#P B3LYP TZVP td=(nstates=2,root=2) nosymm Freq=(HPModes)

Phenoxyl radical - S2 TDDFT - Frequences

 0  2
 O                    -0.00002  -0.13421   0.08764
 C                    -0.00001  -0.05709   1.32916
 C                     1.19565   0.01015   2.13303
 C                     1.17959  -0.00275   3.56649
 C                    -1.19564   0.01046   2.13303
 C                    -1.17959  -0.00243   3.56648
 C                     0.        0.00304   4.29533
 H                     2.14742   0.04605   1.61652
 H                     2.13226  -0.00225   4.08495
 H                    -2.1474    0.04661   1.61651
 H                    -2.13226  -0.00166   4.08494
 H                    0.        -0.01592   5.37493
```

State 2 (step5): Initial state

```
%Chk=phenoxyl_S0.chk
#P b3lyp TZVP FREQ=(HPMODES,VIBROT,READANHARM) nosymm

 Phenoxyl S0 (opt,freq) + reorient + calc FC

 0 2
O         -0.03370        0.13686        0.14749
C         -0.02110        0.06304        1.39847
C          0.88507        0.86725        2.19459
C          0.88779        0.77586        3.56447
C         -0.91205       -0.83030        2.11252
C         -0.88711       -0.90069        3.48342
C          0.00733       -0.10339        4.21826
H          1.55054        1.53449        1.66167
H          1.56882        1.38111        4.15065
H         -1.58891       -1.43099        1.51831
H         -1.55698       -1.57147        4.00791
H          0.01824       -0.16719        5.29925

 NoReord FC CALC1 CHK2 PRTMAT=12
```

F.2 Reoriented/superposed structures

Initial state

```
O   -0.022648    0.132622   -2.247202
C   -0.010048    0.058802   -0.996222
C    0.896122    0.863012   -0.200102
C    0.898842    0.771622    1.169778
C   -0.900998   -0.834538   -0.282172
C   -0.876058   -0.904928    1.088728
C    0.018382   -0.107628    1.823568
H    1.561592    1.530252   -0.733022
H    1.579872    1.376872    1.755958
H   -1.577858   -1.435228   -0.876382
H   -1.545928   -1.575708    1.613218
H    0.029292   -0.171428    2.904558
```

Final state

O	0.005053	0.106083	-2.304993
C	-0.011068	0.063052	-1.061930
C	0.835218	0.868394	-0.216024
C	0.874920	0.733988	1.210718
C	-0.902192	-0.772758	-0.295366
C	-0.839171	-0.885115	1.132432
C	0.035624	-0.129486	1.898227
H	1.486683	1.588735	-0.696273
H	1.582037	1.347008	1.758963
H	-1.633754	-1.358814	-0.838784
H	-1.516408	-1.579721	1.617460
H	0.080659	-0.229093	2.972446

Bibliography

[1] P. Jensen and P. R. Bunker, *Computational Molecular Spectroscopy*, John Wiley and Sons Ltd, Chichester, UK, 2000.

[2] S. Carter, N. C. Handy, C. Puzzarini, R. Tarroni, and P. Palmieri, *Molecular Physics* **98**, 1697 (2000).

[3] M. Biczysko, R. Tarroni, and S. Carter, *The Journal of Chemical Physics* **119**, 4197 (2003).

[4] D. C. Clary, *SCIENCE* **314**, 5797 (2006).

[5] V. Barone and A. Polimeno, *Chemical Society Reviews* **36**, 1724 (2007).

[6] V. Barone, R. Improta, and N. Rega, *Accounts of Chemical Research* **41**, 605 (2008).

[7] H. Koeppel, W. Domcke, and L. S. Cederbaum, *Advances in Chemical Physics* **57**, 59 (1984).

[8] A. Lami, C. Petrongolo, and F. Santoro, *Conical Intersections, Electronic Structure, Dynamics and Spectroscopy*, chapter Absorption, emission and photoelectron continuous-wave spectra, World Scientific Publishing Co., Singapore, 2004.

[9] H.-D. Meyer, U. Manthe, and L. S. Cederbaum, *Chemical Physics Letters* **171**, 97 (1990).

[10] A. Raab, G. Worth, and H.-D. Meyer, *The Journal of Chemical Physics* **110**, 936 (1999).

[11] P. Macak, Y. Luo, and H. Ågren, *Chemical Physics Letters* **330**, 447 (2000).

[12] S. Grimme, *Reviews in Computational Chemistry* **20**, 153 (2004).

[13] G. Scalmani, M. J. Frish, B. Menucci, J. Tomasi, R. Cammi, and V. Barone, *Physical Review Letters* **124**, 094107 (2006).

[14] F. Furche and R. Ahlrichs, *The Journal of Chemical Physics* **121**, 12772 (2004).

[15] A. Köhn and C. Hättig, *The Journal of Chemical Physics* **199**, 5021 (2003).

[16] J. Franck, *Transactions of the Faraday Society* **21**, 536 (1926).

[17] E. Condon, *Physical Review* **32**, 858 (1928).

[18] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. A. Montgomery, Jr., T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson,

H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez, and J. A. Pople, Gaussian development version, Revision G.03, Gaussian, Inc., Wallingford, CT, 2006.

[19] F. Santoro, R. Improta, A. Lami, J. Bloino, and V. Barone, *The Journal of Chemical Physics* **128**, 224311 (2008).

[20] J. Bloino, M. Biczysko, O. Crescenzi, and V. Barone, *The Journal of Chemical Physics* **128**, 244105 (2008).

[21] V. Barone, J. Bloino, M. Biczysko, and F. Santoro, *Journal of Chemical Theory and Computation* (2008), submitted.

[22] A. Sayvetz, *The Journal of Chemical Physics* **7**, 383 (1939).

[23] C. Eckart, *Physical Review* **47**, 552 (1935).

[24] S. Califano, *Vibrational States*, John Wiley & Sons, 1976.

[25] G. Herzberg, *Molecular spectra and molecular structure vol 1: spectra of diatomic molecules*, Van Nostrand, 1950.

[26] P. Dirac, *Proceedings of the Royal Society of London* **114**, 243 (1927).

[27] E. Fermi, *Nuclear Physics*, University Of Chicago Press, 1950.

[28] S. Campbell and C. Meyer, *Generalized Inverses of Linear Transformations*, Dover Publications, 1991.

[29] G. Herzberg and E. Teller, *Zeitschrift für Physikalische Chemie - Abteilung B* **21**, 410 (1933).

[30] P. M. Johnson, H. Xu, and T. J. Sears, *The Journal of Chemical Physics* **125**, 164330 (2006).

[31] F. Duschinsky, *Acta Physicochimica URSS* **7**, 551 (1937).

[32] N. Lucas, *Journal of Physics B* **6**, 155 (1973).

[33] E. Condon, *Physical Review* **28**, 1182 (1926).

[34] E. Hutchisson, *Physical Review* **36**, 410 (1930).

[35] E. Hutchisson, *Physical Review* **37**, 45 (1931).

[36] C. Manneback, *Physica* **17**, 1001 (1951).

[37] W. Smith, *Journal of Physics B* **2**, 1 (1969).

[38] J. Katriel and G. Adam, *Journal of Physics B* **3**, 13 (1970).

[39] J. Katriel, *Journal of Physics B* **3**, 1315 (1970).

[40] D. Craig, *Journal of the Chemical Society* , 2146 (1950).

[41] J. Coon, R. DeWames, and C. Loyd, *Journal of Molecular Spectroscopy* , 285 (1962).

[42] T. Anno and A. Sadô, *The Journal of Chemical Physics* , 1611 (1960).

[43] T. E. Sharp and H. M. Rosenstock, *The Journal of Chemical Physics* **41**, 3453 (1964).

[44] R. Botter, V. Dibeler, J. Walker, and R. H.M., *The Journal of Chemical Physics* **44**, 1271 (1966).

[45] J. Weber and G. Hohlneicher, *Molecular Physics* **101**, 2125 (2003).

[46] J. Wolf and J. Weber, Vibrola (version 2.5), 2002, Universität zu Köln, Germany.

[47] V. I. Baranov, G. L. A., and B. K. Novosadov, *Journal of Molecular Structure* **70**, 1 (1981).

[48] V. I. Baranov and G. L. A., *Journal of Molecular Structure* **70**, 31 (1981).

[49] D. W. Kohn, E. S. J. Robles, C. F. Logan, and P. Chen, *Journal of Physical Chemistry* **97**, 4936 (1993).

[50] A. M. Mebel, Y.-T. Chen, and S.-H. Lin, *Chemical Physics Letters* **258**, 53 (1996).

[51] A. M. Mebel, Y.-T. Chen, and S.-H. Lin, *The Journal of Chemical Physics* **105**, 9007 (1996).

[52] R. Islampour, M. Dehestani, and S. H. Lin, *Journal of Molecular Spectroscopy* **194**, 179 (1999).

[53] H. Kikuchi, M. Kubo, N. Watanabe, and H. Suzuki, *The Journal of Chemical Physics* **119**, 729 (2003).

[54] J.-L. Chang, *Journal of Molecular Spectroscopy* **232**, 102 (2005).

[55] J.-L. Chang, *The Journal of Chemical Physics* **128**, 174111 (2008).

[56] T. R. Faulkner and F. S. Richardson, *The Journal of Chemical Physics* **70**, 1201 (1979).

[57] K. Chen and C. Pei, *Chemical Physics Letters* **165**, 523 (1990).

[58] K. C. Kulander, *The Journal of Chemical Physics* **71**, 2736 (1979).

[59] R. Borrelli and A. Peluso, *The Journal of Chemical Physics* **129**, 064116 (2008).

[60] P. T. Ruhoff, *Chemical Physics* **186**, 355 (1994).

[61] P. T. Ruhoff and M. A. Ratner, *International Journal of Quantum Chemistry* **77**, 383 (2000).

[62] J. Lermé, *Chemical Physics* **145**, 67 (1990).

[63] H. Kupka and C. P. H., *The Journal of Chemical Physics* **85**, 1303 (1986).

[64] E. V. Doktorov, I. A. Malkin, and V. I. Man'ko, *Journal of Molecular Spectroscopy* **56**, 1 (1975).

[65] E. V. Doktorov, I. A. Malkin, and V. I. Man'ko, *Journal of Molecular Spectroscopy* **64**, 302 (1977).

[66] R. J. Glauber, *Physical Review Letters* **10**, 84 (1963).

[67] D. Gruner and P. Brumer, *Chemical Physics Letters* **138**, 310 (1987).

[68] R. Berger, C. Fischer, and M. Klessinger, *Journal of Physical Chemistry A* **102**, 7157 (1998).

[69] A. Hazra and M. Nooijen, *International Journal of Quantum Chemistry* **95**, 643 (2003).

[70] M. Dierksen and S. Grimme, *The Journal of Chemical Physics* **122**, 244101 (2005).

[71] J. Liang and H. Li, *Molecular Physics* **103**, 3337 (2005).

[72] P.-Å. Malmqvist and N. Forsberg, *Chemical Physics* **228**, 227 (1997).

[73] A. Hazra, H. H. Chang, and M. Nooijen, *The Journal of Chemical Physics* **121**, 2125 (2004).

[74] J. M. Luis, D. M. Bishop, and B. Kirtman, *The Journal of Chemical Physics* **120**, 813 (2004).

[75] J. M. Luis, M. Torrent-Succarat, M. Solà, D. M. Bishop, and B. Kirtman, *The Journal of Chemical Physics* **122**, 184104 (2005).

[76] S. Roman, *The Umbral Calculus*, Academic Pr, 1984.

[77] I. Wolfram Research, *Mathematica*, Wolfram Research, Inc., Champaign, Illinois, version 5.2 edition, 2005.

[78] D. E. Knuth, *Art of Computer Programming, Volume 1: Fundamental Algorithms (3rd ed.)*, Addison-Wesley, 1997.

[79] N. Wirth, *Algorithms and Data Structures*, Prentice Hall, 1985.

[80] S. Schumm, M. Gerhards, and K. Kleinermanns, *Journal of Physical Chemistry A* **104**, 10645 (2000).

[81] K. M. Ervin, T. M. Ramond, G. E. Davico, R. L. Schwartz, S. M. Casey, and W. C. Lineberger, *Journal of Physical Chemistry A* **105**, 10822 (2001).

[82] F. Santoro, R. Improta, A. Lami, J. Bloino, and V. Barone, *The Journal of Chemical Physics* **126**, 084509 (2007).

[83] H.-C. Jankowiak, J. L. Stuber, and R. Berger, *The Journal of Chemical Physics* **127**, 234101 (2007).

[84] R. Berger and M. Klessinger, *Journal of Computational Chemistry* **18**, 1312 (1997).

[85] M. J. H. Kemper, J. M. F. Van Dijk, and H. M. Buck, *Chemical Physics Letters* **53**, 121 (1978).

[86] T. Beyer and D. F. Swinehart, *Communications of the Association for Computing Machinery* **16**, 379 (1973).

[87] K. M. Ervin, J. Ho, and W. C. Lineberger, *Journal of Physical Chemistry* **92**, 5405 (1988).

[88] P. Hohenberg and W. Kohn, *Physical Review* **136**, B864 (1964).

[89] W. Kohn and L. J. Sham, *Physical Review* **140**, A1133 (1965).

[90] E. Runge and E. K. U. Gross, *Physical Review Letters* **52**, 997 (1984).

[91] K. Burke, J. Werschnik, and E. K. U. Gross, *The Journal of Chemical Physics* **123**, 062206 (2005).

[92] M. Dierksen and S. Grimme, *The Journal of Chemical Physics* **120**, 3544 (2004).

[93] M. Dierksen and S. Grimme, *Journal of Physical Chemistry A* **108**, 10225 (2004).

[94] F. Santoro, R. Improta, A. Lami, and V. Barone, *The Journal of Chemical Physics* **126**, 184102 (2007).

[95] F. Santoro, $\mathcal{FC}classes$, a fortran 77 code, 2008, visit http://village.ipcf.cnr.it.

[96] L. S. Cederbaum and W. Domcke, *The Journal of Chemical Physics* **64**, 603 (1976).

[97] D. Begue, P. Carbonniere, and C. Pouchan, *Journal of Physical Chemistry A* **105**, 11379 (2001).

[98] S. Tokura, K. Yagi, T. Tsuneda, and K. Hirao, *Chemical Physics Letters* **436**, 30 (2007).

[99] J. M. Luis, B. Kirtman, and O. Christiansen, *The Journal of Chemical Physics* **125**, 154114 (2006).

[100] G. Rauhut and P. Pulay, *Journal of Physical Chemistry* **99**, 3093 (1995).

[101] P. Sinha, S. E. Boesch, C. Gu, R. A. Wheeler, and A. K. Wilson, *Journal of Physical Chemistry A* **108**, 9213 (2004).

[102] M. P. Andersson and P. Uvdal, *Journal of Physical Chemistry A* **109**, 2937 (2005).

[103] V. Barone, *The Journal of Chemical Physics* **122**, 014108 (2005).

[104] J. Spanget-Larsen, M. Gil, A. Gorski, D. M. Blake, J. Waluk, and J. G. Radziszewski, *Journal of the American Chemical Society* **123**, 11253 (2001).

[105] J. G. Radziszewski, M. Gil, A. Gorski, J. Spanget-Larsen, J. Waluk, and B. J. Mróz, *The Journal of Chemical Physics* **115**, 9733 (2001).

[106] R. F. Gunion, M. K. Gilles, M. L. Ploak, and W. C. Lineberg, *International Journal of Molecular Spectroscopy* **117**, 621 (1992).

[107] V. Barone and P. Cimino, *Chemical Physics Letters* **454**, 139 (2008).

[108] J. C. Bopp, J. R. Roscioli, M. A. Johnson, T. M. Miller, A. A. Viggiano, S. M. Villano, W. S. W., and L. W. C., *Journal of Physical Chemistry A* **111**, 1214 (2007).

[109] R. Borrelli, *Chemical Physics Letters* **445**, 84 (2007).

[110] C. Moeller and M. S. Plesset, *Physical Review* **46**, 618 (1934).

[111] T. H. Dunning, *The Journal of Chemical Physics* **90**, 1007 (1989).

[112] R. Kendall, T. Dunning Jr., and R. Harrison, *The Journal of Chemical Physics* **96**, 6769 (1992).

[113] C. Hampel, K. Peterson, and H.-J. Werner, *Chemical Physics Letters* **190**, 1 (1992).

[114] M. J. O. Deegan and P. J. Knowles, *Chemical Physics Letters* **227**, 321 (1994).

[115] T. Korona and H.-J. Werner, *The Journal of Chemical Physics* **118**, 300 (2003).

[116] H.-J. Werner, P. J. Knowles, R. Lindh, F. R. Manby, M. Schütz, P. Celani, T. Korona, G. Rauhut, R. D. Amos, A. Bernhardsson, A. Berning, D. L. Cooper, M. J. O. Deegan, A. J. Dobbyn, F. Eckert, C. Hampel, G. Hetzer, A. W. Lloyd, S. J. McNicholas, W. Meyer, M. E. Mura, A. Nicklass, P. Palmieri, R. Pitzer, U. Schumann, H. Stoll, A. J. Stone, R. Tarroni, and T. Thorsteinsson, Molpro, version 2006.1, a package of ab initio programs, 2006, see http://www.molpro.net.

[117] E. P. Grimsrud, S. Chowdhury, and P. Kebarle, *The Journal of Chemical Physics* **83**, 1059 (1985).

[118] S. Vassiliev and D. Bruce, *Photosynthesis Research* **97**, 75 (2008).

[119] T. Noguchi, *Plant and Cell Physiology* **43**, 1112 (2002).

[120] C. Adamo and V. Barone, *The Journal of Chemical Physics* **110**, 6158 (1999).

[121] G. Brancato, N. Rega, and V. Barone, *The Journal of Chemical Physics* **125**, 164515 (2006).

[122] M. Cossi, G. Scalmani, N. Rega, and V. Barone, *Journal of Computational Chemistry* **24**, 669 (2003).

[123] K. W. Paulisse, T. O. Friday, M. L. Graske, and W. F. Polik, *The Journal of Chemical Physics* **113**, 184 (2000).

[124] T. Shimanouchi, *Tables of Molecular Vibrational Frequencies Consolidated, Vol. I*, National Bureau of Standards, Washington D.C., 1972.

[125] T. Vreven and K. Morokuma, *Journal of Computational Chemistry* **21**, 1419 (2000).

[126] A. K. Rappè, C. J. Casewit, K. S. Colwell, W. A. Goddard III, and W. M. Skiff, *Journal of the American Chemical Society* **114**, 10024 (1992).

[127] A. B. Trofimov, J. Schirmer, V. B. Kobychev, A. W. Potts, D. M. P. Holland, and L. Karlsson, *Journal of Physics B* **39**, 305 (2006).

[128] M. Gouterman, *Journal of Molecular Spectroscopy* **6**, 138 (1961).

[129] B. Minaev, Y.-H. Wang, G.-K. Wang, Y. Luo, and H. Ågren, *Spectrochimica Acta, Part A* **65**, 308 (2006).

[130] G. Hohlneicher and J. Wolf, *Berichte der Bunsengesellschaft für Physikalische Chemie* **99**, 366 (1995).

[131] K. G. Geigle, J. Wolf, and G. Hohlneicher, *Journal of Photochemistry and Photobiology, A* **105**, 183 (1997).

[132] K. G. Geigle and G. Hohlneicher, *Journal of Photochemistry and Photobiology, A* **247**, 480 (1999).

[133] D. P. Craig and G. J. Small, *The Journal of Chemical Physics* **50**, 3827 (1969).

[134] A. T. Gradyushko, K. N. Solovev, and A. S. Starukhin, *Optics and Spectroscopy* **40**, 267 (1976).

[135] A. T. Gradyushko, K. N. Solovev, A. S. Starukhin, and A. M. Shulga, *Optics and Spectroscopy* **43**, 37 (1977).

[136] J. G. Radziszewsky, J. Waluk, M. Nepraš, and J. Michl, *Journal of Physical Chemistry* **95**, 1963 (1991).

[137] D. Reha, H. Valdes, J. Vondrasek, P. Hobza, A. Abu-Riziq, B. Crews, and M. S. de Vries, *Chemistry - A European Journal* **11**, 6803 (2005).

[138] A. Musgrave and T. G. Wright, *The Journal of Chemical Physics* **122**, 074312 (2005).

[139] M. Schmitt, U. Henrichs, and K. Kleinermanns, *The Journal of Chemical Physics* **103**, 9918 (1995).

[140] E. Nir, K. Kleinermanns, and M. S. de Vries, *Nature* **408**, 949 (2000).

[141] L. C. Snoek, R. T. Kroemer, M. R. Hockridge, and J. P. Simons, *Physical Chemistry Chemical Physics* **3**, 1819 (2001).

[142] E. Arunan and H. S. Gutowsky, *The Journal of Chemical Physics* **98**, 4294 (1993).

[143] J. R. Roscioli and D. W. Pratt, *Proceedings of the National Academy of Sciences* **100**, 13572 (2003).

[144] M. Schmitt, M. Bohm, C. Ratzer, D. Krugler, K. Kleinermanns, I. Kalkman, G. Berden, and W. L. Meerts, *ChemPhysChem* **7**, 1241 (2006).

[145] G. Berden, W. Meerts, M. Schmitt, and K. Kleinermanns, *The Journal of Chemical Physics* **104**, 972 (1996).

[146] M. Becucci, G. Pietraperzia, M. Pasquini, G. Piani, A. Zoppi, R. Chelli, E. Castellucci, and W. Demtroeder, *The Journal of Chemical Physics* **120**, 5601 (2004).

[147] B. M. Giuliano and W. Caminati, *Angewandte Chemie International Edition* **44**, 603 (2004).

[148] M. Schmitt, M. Bhom, C. Ratzer, C. Vu, I. Kalkman, and W. L. Meerts, *Journal of the American Chemical Society* **127**, 10356 (2005).

[149] J. R. Clarkson, E. Baquero, V. A. Shubert, E. M. Myshakin, K. J. Jordan, and T. S. Zwier, *Science* **307**, 1143 (2005).

[150] M. Schmitt, C. Ratzer, and W. Meerts, *The Journal of Chemical Physics* **120**, 2752 (2004).

[151] G. Piani, M. Pasquini, G. Pietraperzia, M. Becucci, N. Schiccheri, M. Biczysko, M. Pavone, and V. Barone, *Journal of Physical Chemistry A* **111**, 12363 (2007).

[152] M. Biczysko, G. Piani, M. Pasquini, N. Schiccheri, G. Pietraperzia, M. Becucci, M. Pavone, and V. Barone, *The Journal of Chemical Physics* **127**, 144303 (2007).

[153] G. Piani, M. Pasquini, G. Pietraperzia, M. Becucci, A. Armentano, and E. Castellucci, *Chemical Physics Letters* **434**, 25 (2007).

[154] R. E. Stratmann, G. E. Scuseria, and M. J. Frisch, *The Journal of Chemical Physics* **109**, 8218 (1998).

[155] O. A. Vydrov and G. E. Scuseria, *The Journal of Chemical Physics* **125**, 234109 (2006).

[156] R. Valero, R. Costa, I. d. P. R. Moreira, D. G. Truhlar, and F. Illas, *The Journal of Chemical Physics* **128**, 114103 (2008).

[157] C. G. Eisenhardt, G. Pietraperzia, and M. Becucci, *Physical Chemistry Chemical Physics* **3**, 1407 (2001).

[158] C. G. Eisenhardt, A. S. Gemechu, H. Baumgärtel, R. Chelli, G. Cardini, and S. Califano, *Physical Chemistry Chemical Physics* **3**, 5358 (2001).

[159] R. Matsumoto, K. Sakeda, Y. Matsushita, T. Suzuki, and T. Ichimuea, *Journal of Molecular Structure* **735-736**, 153 (2005).

[160] L. J. H. Hoffmann, S. Marquardt, A. S. Gemechu, and H. Baumgärtel, *Physical Chemistry Chemical Physics* **8**, 2360 (2006).

[161] W. J. Balfour, *Spectrochim. Acta, Part A* **39**, 795 (1983).

[162] E. B. Wilson, *Physical Review* **45**, 706 (1934).

[163] G. Varsanyi, *Vibrational Spectra of Benzene Derivatives*, Academic Press, New York, 1969.

[164] L. S. Cederbaum and W. Domcke, *Advances in Chemical Physics* **36**, 205 (1977).

[165] L. S. Cederbaum, H. Köppel, and W. von Niessen, *Molecular Physics* **34**, 1759 (1977).

[166] W. J. Balfour, *Journal of Molecular Spectroscopy* **109**, 60 (1985).

[167] S. Tsuzuki, H. Houjou, Y. Nagawa, and H. K., *Journal of Physical Chemistry A* **104**, 1332 (2000).

[168] F. De Angelis, F. Santoro, M. K. Nazeruddin, and V. Barone, *Journal of Physical Chemistry B* **112**, 13181 (2008).

List of publications related to this thesis

- F. Santoro, R. Improta, A. Lami, J. Bloino, V. Barone, *Effective method to compute Franck-Condon integrals for optical spectra of large molecules in solution*, Journal of Chemical Physics **126**, 084509 (2007)
- F. Santoro, A. Lami, R. Improta, J. Bloino, V. Barone, *Effective method for the computation of optical spectra of large molecules at finite temperature including the Duschinsky and Herzberg-Teller effect: The Q_x band of porphyrin as a case study*, Journal of Chemical Physics **128**, 224311 (2008)
- J. Bloino, M. Biczysko, O. Crescenzi, V. Barone, *Integrated computational approach to vibrationally resolved electronic spectra: Anisole as a test case*, Journal of Chemical Physics **128**, 244105 (2008)
- V. Barone, J. Bloino, M. Biczysko, F. Santoro, *Fully integrated approach to compute vibrationally resolved optical spectra: from small molecules to macrosystem*, Journal of Chemical Theory and Computation, *Submitted*